



UNIVERSIDAD NACIONAL DE COLOMBIA

Método para la integración de datos de múltiples sensores de flujo en líneas de conducción de hidrocarburos refinados

Ricardo Hernández Herrera

Universidad Nacional de Colombia

Facultad de Ingeniería, Departamento de Ingeniería Eléctrica y Electrónica

Bogotá, Colombia

2014

Método para la integración de datos de múltiples sensores de flujo en líneas de conducción de hidrocarburos refinados

Ricardo Hernández Herrera

Tesis presentada como requisito parcial para optar al título de:
Magister en Ingeniería Automatización Industrial

Director:

M.Sc. Carlos Eduardo Sánchez

Línea de Investigación:

Automatización de Procesos

Universidad Nacional de Colombia

Facultad de Ingeniería, Departamento de Ingeniería Eléctrica y Electrónica

Bogotá DC, Colombia

2014

La ciencia por sí misma no podrá defender los efectos de las acciones del ser humano, pero ella podrá ciertamente iluminar las posibles consecuencias de los efectos alternativos de las acciones.

Carl Sagan

Agradecimientos

Un especial reconocimiento a la Universidad Nacional de Colombia por haber construido el espíritu investigativo a través de estos años de aprendizaje, no solo en el curso de la maestría, sino también en el pregrado que curse en la sede de Manizales.

Adicionalmente deseo agradecer a mis maestros en automatización, los ingenieros Álvaro Wills Herrera y Álvaro Hernández Mora, tanto en los conocimientos básicos y esenciales de esta área de la ingeniería como en los aspectos profesionales que conllevan.

Resumen

Recientemente en Colombia, en los sistemas de transporte por ducto de hidrocarburos se han instalado medidores que reportan el estado de la operación de transporte sin un mayor análisis de la información reportada. Este trabajo presenta una estrategia metodológica para la fusión de la información de los medidores de flujo instalados a lo largo de los ductos enfocado a la identificación, aislamiento y caracterización de pérdidas o ganancias volumétricas en los diferentes segmentos limitados por estos sensores. Adicionalmente, se evaluarán diferentes condiciones operativas probables así como fallas operativas comunes en función de la operatividad de la metodología propuesta.

Palabras clave: Fusión de información, Fusión Dempster-Shafer, Fusión Bayes, Sistemas de transporte por ducto.

Abstract

Recently in Colombia, at pipeline transportation systems have been installed meters that report the status of the transport operation without further analysis of the information reported. This paper presents a methodological strategy for merging the information flow meters installed along the pipelines focused on the identification, isolation and characterization of volumetric gains and losses in different segments bound by these sensors. Additionally, we evaluate different operating conditions and operational failures likely common function of the operability of the proposed methodology.

Keywords: Information fusion, Dempster-Shafer fusion, Bayes fusion, Pipelines.

Contenido

	Pág.
Resumen.....	VII
Lista de figuras.....	XI
Lista de tablas	XIII
Lista de Símbolos y abreviaturas.....	XIV
Introducción	1
1. Estado del arte.....	3
1.1 Problemática actual y trabajo relacionados.....	3
1.2 Motivación	4
1.3 Objetivos.....	5
1.4 Estructura de la tesis	6
2. Marco Teórico.....	7
2.1 Medición en el transporte de hidrocarburos por ductos.....	8
2.1.1 Ecuación de pérdida/ganancia (L/G)	8
2.1.2 Solución de desviaciones.....	9
2.1.3 Imprecisiones e incertidumbres.....	10
2.1.4 Explicación de L/G	13
2.2 Sistemas de medición de hidrocarburos comunes en ductos.....	14
2.2.1 Medidores de turbina.....	15
2.2.2 Medidores ultrasónicos	19
2.3 Incertidumbre.....	22
2.3.1 Consideraciones prácticas	24
2.3.2 Modelamiento de la medición.....	26
2.4 Métodos para la fusión de la información.....	33
2.4.1 Filtro Kalman.....	34
2.4.2 Fusión de la información usando filtro Kalman (KL)	37
2.4.3 Fusión Bayes para datos simples (BY).....	37
2.4.4 Fusión Dempster-Shafer para datos simples (DS)	38
2.4.5 Fusión por promedio ponderado de confianza	40
2.5 Definición de PDF para mediciones	41
2.5.1 Método Monte Carlo (MCM).....	42
2.6 Teorema de la desigualdad de Chebyshev	45
3. Metodología	47

3.1	Arquitectura conceptual.....	47
3.2	Arquitectura normalizada de un DTHR instrumentado.....	50
3.2.1	Modelo de operación	51
3.2.2	Modelo de desviación operativa por pérdida o ganancia (L/G).....	52
3.3	Niveles metodológicos a implementar	53
3.3.1	Filtro de la información (nivel 1)	54
3.3.2	Identificación de desviaciones operativas L/G (nivel 2)	56
3.3.3	Aislamiento de desviaciones L/G en segmentos <i>mij</i> (nivel 3)	58
3.3.4	Caracterización de una desviación operativa L/G (nivel 4).....	61
4.	Evaluación Teórica	65
4.1	Probabilidad de mediciones de medidores con diferente incertidumbre.	66
4.2	Evaluación de filtros de información (nivel 1).....	69
4.2.1	Evaluación de Filtro Kalman para mediciones Interlínea	69
4.2.2	Evaluación de Filtro Ponderado para mediciones Interlínea.....	74
4.2.3	Filtros Kalman y Ponderado usando desigualdad de Chebyshev.....	77
4.3	Evaluación de identificación de L/G por medidores <i>mi</i> (nivel 2)	78
4.3.1	Acondicionamiento de observaciones hechas por medidor <i>mi</i>	79
4.3.2	Evaluación de las mediciones de <i>mi</i> cuando existe una L/G posterior	80
4.3.3	Evaluación de las mediciones de <i>mi</i> cuando existe una L/G anterior	84
4.3.4	Evaluación de sensibilidad de identificación de pérdidas/ganancias	86
4.4	Evaluación de análisis de aislamiento de L/G para segmentos <i>mij</i> (nivel 3)	89
4.4.1	Evaluación de aislamiento de L/G en segmentos <i>mij</i> por BY	91
4.4.2	Evaluación de aislamiento de L/G en segmentos <i>mij</i> por DS.....	100
4.5	Caracterización y cuantificación de desviaciones L/G (nivel 4).....	108
5.	Otras condiciones operativas	113
5.1	Desconexión de medidor de flujo	114
5.2	Descalibración del medidor de flujo	116
5.3	Congelamiento de la señal del medidor de flujo	119
5.4	Desincronización de información de los medidores	121
6.	Conclusiones y recomendaciones	125
6.1	Conclusiones.....	125
6.2	Recomendaciones.....	127
	Bibliografía	129
A.	Anexo: Simulación en Condiciones de Operación Reales.....	133
B.	Anexo: Umbrales de Sensibilidad.....	135

Lista de figuras

	Pág.
Figura 2-1: Partes de un medidor de turbina [10].	15
Figura 2-2: Rendimiento de medidores de turbina entre factores M y K [3].	16
Figura 2-3: Esquema de montaje de transductores ultrasónicos [11].	19
Figura 2-4: Modelo de medición del tiempo transitorio entre los transductores [9].	21
Figura 2-5: Proceso MCM para cuantificar PDF para medición [7].	45
Figura 3-1: Modelo JDL para fusión de datos [16].	49
Figura 3-2: Modelo para la integración de datos de un DTHR.	50
Figura 3-3: Sistema simplificado de transporte de hidrocarburos por ductos.	51
Figura 3-4: Modelo de pérdida/ganancia para el análisis del segmento <i>mij</i> .	53
Figura 3-5: Análisis de probabilidad para los medidores del segmento <i>mij</i> .	57
Figura 3-6: Análisis de probabilidad para observaciones de los medidores T.	59
Figura 4-1: Diferencia de PDF entre medidores terminales e interlínea.	66
Figura 4-2: Probabilidad de mediciones entre un medidor interlínea y terminal	67
Figura 4-3: Análisis de probabilidad para los medidores del segmento <i>mij</i> .	68
Figura 4-4: Filtro Kalman para mediciones de flujo Interlínea.	70
Figura 4-5: Filtro Kalman para mediciones de flujo Interlínea a bajas frecuencias.	72
Figura 4-6: Filtro Kalman para mediciones de flujo Interlínea a variaciones rápidas.	73
Figura 4-7: Filtro Ponderado para mediciones de flujo.	74
Figura 4-8: Filtro Ponderado para mediciones de flujo a bajas frecuencias.	76
Figura 4-9: Filtro Ponderado para mediciones de flujo a variaciones rápidas.	77
Figura 4-10: Respuesta de intersección entre medidores terminales e interlínea.	78
Figura 4-11: Modelo simplificado para la medición de probabilidad de L/G posterior.	81
Figura 4-12: Valor de la probabilidad con y sin L/G posterior al medidor <i>mint</i> .	82
Figura 4-13: Diferencia de probabilidades de la figura 4.12.	82
Figura 4-14: Modelo simplificado para la medición de probabilidad de L/G anterior.	84
Figura 4-15: Valor de la probabilidad con y sin L/G posterior al medidor <i>mint</i> .	85

Figura 4-16:	Diferencia de probabilidades de la figura 4.15.	85
Figura 4-17:	Valor de la probabilidad para pruebas de sensibilidad.	87
Figura 4-18:	Diferencia de la probabilidad para pruebas de sensibilidad.	88
Figura 4-19:	Método de fusión de información partiendo de medidores terminales. ...	90
Figura 4-20:	Método de fusión de información partiendo de medidores Interlínea.	90
Figura 4-21:	Fusión de datos con referencia terminales con L/G 1% por BY.	91
Figura 4-22:	Fusión de datos con referencia interlínea con L/G 1% por BY.	93
Figura 4-23:	Fusión para un segmento <i>mij</i> por BY con L/G del 1% al 0,1% T.	94
Figura 4-24:	Fusión para un segmento <i>mij</i> por BY con L/G del 1% al 0,1% I.	95
Figura 4-25:	Fusión datos múltiples segmentos <i>mij</i> por BY, única L/G de 1% T.	96
Figura 4-26:	Fusión datos múltiples segmentos <i>mij</i> por BY, única L/G de 1% I.	98
Figura 4-27:	Fusión datos múltiples segmentos <i>mij</i> por BY, múltiples L/G de 1% T...	99
Figura 4-28:	Fusión de datos con referencia terminales con L/G 1% por DS.	101
Figura 4-29:	Fusión de datos con referencia interlínea con L/G 1% por DS.	103
Figura 4-30:	Fusión para un segmento <i>mij</i> por DS con L/G del 1% al 0,1% T.	104
Figura 4-31:	Fusión para un segmento <i>mij</i> por DS con L/G del 1% al 0,1% I.	105
Figura 4-32:	Fusión datos múltiples segmentos <i>mij</i> por DS, única L/G de 1% T.	106
Figura 4-33:	Fusión datos múltiples segmentos <i>mij</i> por DS, única L/G de 1% I.	107
Figura 4-34:	Fusión datos múltiples segmentos <i>mij</i> por DS, múltiples L/G de 1% T.	108
Figura 5-1:	Falla de desconexión con referencia medidores terminales	114
Figura 5-2:	Falla de desconexión con referencia medidores interlínea.	115
Figura 5-3:	Falla de descalibración con referencia medidores terminales.	117
Figura 5-4:	Falla de descalibración con referencia medidores interlínea.	118
Figura 5-5:	Falla por congelamiento con referencia medidores terminales.	120
Figura 5-6:	Falla por congelamiento con referencia medidores interlínea.	121
Figura 5-7:	Falla por desincronización <i>m4</i> y <i>m5</i> con referencia terminales.	122
Figura 5-8:	Falla por desincronización <i>m4</i> y <i>m5</i> con referencia interlínea.	123
Figura B-1:	Definición de umbral con Dempster- Shafer y Bayes.	135
Figura B-2:	Distribución valores para el umbral cero con DS y BY.	136
Figura B-3:	Segmentación para análisis de umbrales de sensibilidad.	137
Figura B-4:	Resultados de análisis probabilístico de métodos.	139

Lista de tablas

	Pág.
Tabla 3-1: Proceso de filtro Kalman.	55
Tabla 3-2: Proceso de filtro de peso ponderado.	56
Tabla 3-3: Proceso de análisis de probabilidad de intersección para medidores.	58
Tabla 3-4: Fusión de datos por segmentos mij por Dempster-Shafer.	60
Tabla 3-5: Fusión de datos por segmentos mij por Bayes.	61
Tabla 3-6: Fusión de datos por segmentos mij por promedio ponderados.	63
Tabla 4-1: Filtro Kalman para varios factores de varianza q	70
Tabla 4-2: Filtro Kalman para varios factores de varianza q a bajas frecuencias.	71
Tabla 4-3: Filtro Kalman para varios factores de varianza q a variaciones rápidas	73
Tabla 4-4: Filtro Ponderado a varios factores de A.	74
Tabla 4-5: Filtro Ponderado a varios factores de A para bajas frecuencias.	75
Tabla 4-6: Filtro Ponderado a varios factores de A para variaciones rápidas.	76
Tabla 4-7: Características estadísticas para medidores a evaluar en Identificación. ..	79
Tabla 4-8: Tabulación de datos de figura 4.27 según ecuación 4.2.	99
Tabla 4-9: Caracterización de L/G de pérdida de 1% en el segmento 4&5.	109
Tabla 4-10: Caracterización de L/G de ganancia de 1% en el segmento 4&5.	110

Lista de Símbolos y abreviaturas

Símbolos con letras latinas

Símbolo	Término	Unidad SI	Definición
k	Momento k	1	
K	Grupo de momentos k	1	
m_i	I-esímo medidor de flujo	1	
M_i	Mediciones del I-esímo medidor de flujo	1	
s	Incertidumbre de variables aleatorias	1	
s^2	Varianza de variables aleatorias	1	
u	Incertidumbre estándar	1	
u_c	Incertidumbre estándar combinada	1	
u^2	Varianza estándar estimada	1	
U	Incertidumbre expandida	1	
x_i	Mediciones estimadas	1	
X_i	Cantidades de entrada	1	
y	Estimado de Medición	1	
Y	Cantidad medida	1	

Símbolos con letras griegas

Símbolo	Término	Unidad SI	Definición
σ	Desviación estándar	1	
μ_q	Valor esperado	1	

Abreviaturas

Abreviatura	Término
BY	Fusión Bayes

Abreviatura Término

<i>DTHR</i>	Ducto de transporte de Hidrocarburos Refinados
<i>KL</i>	Filtro Kalman
<i>L/G</i>	Pérdida o Ganancia.
<i>PDF</i>	Función de densidad de probabilidad
<i>SI</i>	Sistema Internacional de medidas

Introducción

Las pérdidas y ganancias en sistemas de transporte son caracterizadas como lícitas e ilícitas. Las pérdidas lícitas van desde roturas de las líneas de producción, quemas operativas, drenajes y evaporaciones de productos almacenados. Las pérdidas ilícitas son generadas por actores que usan este vehículo para financiar otras actividades delictivas o para lucrarse directamente de la venta de estos a un costo inferior al del mercado.

En el país y en el mundo, Ecopetrol es líder en la reducción de pérdidas por acción ilícita en líneas de transporte de hidrocarburos, generando estrategias, en gran medida policivas, con el fin de asegurar el transporte de combustibles a lo largo del país. En el 2002 el apoderamiento de hidrocarburos llegó a su máximo histórico de 7270 barriles día y en la actualidad se ha llegado a niveles de 30 barriles día, lo que infiere una reducción del 99%.

El producto hurtado de los sistemas de transporte es extraído por grupos organizados al margen de la ley. Estos se lucran de forma sistemática, generando pérdidas económicas y ambientales en las zonas donde ocurren estos eventos. Los productos son extraídos de utilizando válvulas ilícitas y lícitas aprovechando la extensión de los sistemas de transporte, poco control de las autoridades y las limitaciones técnicas inherentes a la operación. En la actualidad se utilizan patrullajes periódicos a lo largo de las líneas de conducción para ubicar estas válvulas o punto de extracción de combustible, es decir, una actividad más reactiva que activa o preventiva.

Las acciones ilícitas ocasionan impactos la integridad y operatividad de los poliductos debido al debilitamiento y fallas mecánicas inducidas al ducto, incrementando el riesgo operación ductos en los linderos donde este se ubica geográficamente. En los accidentes presentados donde la no detección oportuna de estas instalaciones ha generado cuantiosos daños ambientales y afectación a población civil desde heridas leves hasta muertes. Dada la potencialidad y la probabilidad de la masificación del riesgo por una

desviación operativa de pérdida o ganancia a comunidades colindantes a los ductos no detectada con oportunidad y certeza, es necesario generar estrategias que vinculen diferentes fuentes, información y supervisión de los ductos adheridos a la operación del mismo.

Es imperativa la necesidad de mejorar los esquemas de detección de este tipo de desviaciones. Debido a las implicaciones económicas y operativas asociadas a la inclusión de nuevas tecnologías de detección, las soluciones planteadas deben buscar aprovechar recursos existentes instalados y operados buscando siempre la mejor alternativa para los operadores de los ductos y la detección de pérdidas de menor rata de flujo. Estas condiciones implican retos importantes a nivel metodológico y de implementación para toda estrategia analizada y sugerida.

A nivel Colombia, las pérdidas o ganancias operativas en transporte de hidrocarburos refinados tienen un impacto fiscal doble para las empresas productoras y transportadas, estas deben reportar su producción a la Agencia Nacional de Hidrocarburos (ANH). Dado que el producto perdido es contabilizado, se deben pagar aranceles y costos operativos asociados a él. El producto sobre-cuantificado fiscalmente debe pagar impuesto de ganancias ocasionales, es decir, un impacto aún más grande para las empresas. Esto implica que las desviaciones operativas deben ser cuantificadas de la mejor manera posible, reduciendo así los impactos fiscales del producto a comercializar.

1.Estado del arte

1.1 Problemática actual y trabajo relacionados

Los esquemas de medición actuales no permiten una adecuada interpretación de las desviaciones de pérdida o ganancia en tiempo real. Estos esquemas parten de la instalación de equipos de baja incertidumbre y alta confiabilidad en los puntos terminales de la línea, el total cuantificado por los dos medidores generan la información de ganancia o pérdida del sistema. Este esquema no permite identificar, aislar o caracterizar la desviación en cualquier segmento de la línea con el agravante que solo se puede detectar estas desviaciones en periodos largos de tiempo y con magnitudes de flujo altas.

Existen sistemas de detección de desviaciones focalizados que parten de modelos hidrodinámicos de la línea [1]. Estos interpretan teóricamente el movimiento del fluido en diferentes sectores del ducto, cuando existe una desviación de cualquier tipo, el sistema estima donde se encuentra esta. Dado que estos sistemas se basan en modelos no configurados para la realidad de la operación actual, la precisión y oportunidad de aislamiento de la desviación se ve afectada por variables como la temperatura ambiente, el desconocimiento real de los volúmenes de almacenamiento de la línea, entre otras [2]. Esta imprecisión generada por estas variables puede llevar a que el sistema sea completamente inútil para las necesidades reales de la operación.

Otros sistemas de detección se basan en la detección de caídas de presión en diferentes sectores de la línea [3] [4]. Se basan en modelos empíricos del cómo se reflejan las caídas de presión y sus tendencias para su posterior interpretación. Estos sistemas son fiables pero dado que parten de la variable presión, pueden ser ocultadas por oscilaciones de los sistemas de control de las bombas de transporte, la temperatura del producto y los cambios de esta a lo largo del ducto, entregas al paso, entre otras. Este sistema puede ser más fiable que el anterior pero con la necesidad que se debe tener un concepto amplio de observabilidad del modelo del sistema para llevarlo a niveles de precisión aceptables para la operación.

Recientemente se han adoptado esquemas de medición de flujo en diferentes puntos sobre la línea con equipos de medición de alta incertidumbre [5]. La incertidumbre es inherente a la medición de variables físicas y se expresa como una proporción de la cantidad medida por el equipo asociado, en este caso en particular, medidores ultrasónicos con una incertidumbre en banco de 3% y en operación mayor al 5%. Dado esto y en función de la cantidad de productos refinados que se mueven actualmente por los sistemas de transporte, los equipos de medición que se encuentran en la línea no pueden detectar pequeñas pérdidas o ganancias del sistema, haciéndolos “insensibles” a dichas desviaciones.

1.2 Motivación

Esta motivación parte de la formulación clásica propuesta por estándar API 2560 [6] presenta fallas importantes al momento de realizar una detección de pérdidas en tiempo real utilizando sensores de flujos dinámicos dado que se deben considerar factores tanto físicos como temporales de las muestras para ser comparadas. La visión que se planteará permitirá anexar un análisis adaptable a las condiciones previas de las observaciones hechas contrastando la condición actual en tiempo real.

Adicional a los problemas subsecuentes con la identificación de pérdidas, la cuantificación lleva una connotación especial dado que los productos transportados cuentan con diferentes características fisicoquímicas que incrementan la dificultad de una cuantificación confiable. Cualquier producto hidrocarburo refinado tiene tolerancias de producción debido a la calidad del petróleo crudo con él se realiza el proceso. Los factores asociados a la producción de refinados inciden directamente en una correcta estimación de cantidades de producto degradando la confiabilidad de la información consolidada, y por ende, la necesidad de un sistema de detección de desviaciones adaptable a los productos medidos.

Aumentar la oportunidad, certeza y detección de desviaciones operativas de los productos transportados puede evitar una falla mayor con pérdidas de vidas humanas al activar medidas de control automáticas y de personal de emergencias. El caso sucedido en Dosquebradas-Risaralda en diciembre de 2010, donde por un evento netamente geológico ocasiono una ruptura total del Poliducto de 6" de Puerto Salgar –Cartago y

como consecuencia se perdieron 33 vidas humanas por una respuesta no oportuna en la activación de los mecanismos de control del ducto y socorro por parte de las autoridades. Esta clase de eventos pueden ser detectados con mayor precisión y mejorar la respuesta de identificación de pérdidas más certeras.

Se propondrá entonces utilizar la información de flujo existente en la operación de un ducto segmentado por medidores de flujo terminales (baja incertidumbre) e interlínea (alta incertidumbre) para el análisis de desviaciones de flujo entre los segmentos. Se mezclará las mediciones hechas por los medidores terminales con los interlínea, utilizando procesos de filtrado y fusión de información utilizados en áreas diferentes a la de la industria del petróleo [7]. Este análisis permitirá identificar, aislar y caracterizar las desviaciones desde un punto de vista holístico a las variables del sistema de transporte con líneas segmentadas por dichos medidores.

1.3 Objetivos

Proponer una metodología conceptual que permita la integración de información de combustibles refinados líquidos en líneas de transporte de hidrocarburos para mejorar la certeza de múltiples medidores de flujo dinámico de alta incertidumbre (FIT) instalados sobre ella.

Se tendrán los siguientes objetivos específicos:

1. Proponer una metodología teórica para la identificación, aislamiento y caracterización de desviaciones de pérdidas o ganancias en líneas de transporte de hidrocarburos refinados instrumentados con múltiples medidores de flujo a lo largo de esta.
2. Evaluar la implementación de la metodología propuesta con dos algoritmos de fusión de información que se ajusten a las condiciones de múltiples mediciones flujo instantáneo en un ducto de transporte de hidrocarburos refinados.
3. Evaluar bajo criterios identificación, aislamiento y caracterización de medición la metodología propuesta con una simulación de condiciones operativas de una línea de transporte de hidrocarburos.

1.4 Estructura de la tesis

Esta tesis se estructurará de la siguiente manera:

En el capítulo 2 se presentará la base teórica en la cual se fundamentará la metodología a proponer. Se iniciará por los esquemas actuales de conciliación de volúmenes transportados en un ductos, se pasará a los equipos de medición utilizados para cuantificar los volúmenes, se presentarán los esquemas clásicos de incertidumbre para finalizar en con algoritmos para la fusión de información y herramientas para el modelamiento de la incertidumbre en sistemas.

Posteriormente, se presentará en el capítulo 3 la propuesta teórica de la metodología a analizar posteriormente. Se iniciará por la presentación del modelo a seguir para el análisis de información de múltiples datos y los modelos de pérdida o ganancia. Se presentará posteriormente los niveles de análisis y las herramientas a utilizar para cada uno de ellas. Se finalizará con las propuestas de cuantificación de pérdidas o ganancias para los segmentos.

En el capítulo 4 se evaluará desde el nivel 1 hasta el nivel 4 de la metodología propuesta en el capítulo 3. Se presentarán varias condiciones viables en la operación normal de un ducto para ser analizada por los diferentes niveles de la metodología propuesta. Esta evaluación determinará las limitaciones teóricas de la metodología.

En el capítulo 5 se analizará como las fallas comunes de los equipos de medición afectarían el desempeño de la metodología a implementar. Este análisis se hará de una manera superficial y solo presentará una posible afectación.

Finalmente, se presentarán las recomendaciones para implementación, posibles áreas de investigación posterior y conclusiones de la metodología propuesta y analizada en esta tesis.

2.Marco Teórico

Dada la cantidad de herramientas que se han utilizado con los métodos de fusión de mediciones y su aplicabilidad en otras industrias diferentes a la del transporte de hidrocarburos refinados, se procederá a dar un marco teórico donde se basará esta investigación.

Esta sección se iniciará con una breve descripción de como la medición en el transporte de hidrocarburos es susceptible de cambiar por factores internos y externos a la misma. Posteriormente, se adentrará en la conceptualización de lo referido a la incertidumbre, pasando por las teorías con las cuales se pueden determinar según guías de organizaciones internacionales. Partiendo de estas dos aproximaciones, se crea la base para aumentar la certeza de detección de desviaciones en los ductos.

Posteriormente se presentarán diferentes alternativas matemáticas para el filtrado y fusión de información, herramientas importantes para el mejoramiento de las características de detección de varios sensores operando simultáneamente en un ducto.

Se finalizaría este marco con análisis de probabilidad de las diferentes mediciones y observaciones que se harán y analizarán por esta metodología. Este análisis es necesario para conservar la integridad estadística de los datos a medida que los diferentes procesos que se propondrán los adecue a sus necesidades sin afectar las propiedades de los datos iniciales.

Todas las herramientas que se mostrarán tendrán la finalidad de aportar a la integración de la información presentada por los sistemas de transporte de hidrocarburos refinados instrumentados con la meta particular de mejorar la información obtenida de esos datos en procura de mejorar la certeza en la detección de desviaciones operativas.

2.1 Medición en el transporte de hidrocarburos por ductos

El transporte por ductos de hidrocarburos es un método ampliamente usado por las compañías petroleras dado sus reducidos costos operacionales con otros métodos de transporte y su operación segura. Este método es riesgoso para los diferentes grupos de interés que se lucran o benefician de dicho transporte. El transporte de hidrocarburos por ductos tiene consideraciones importantes y críticas que el transporte de otros productos, por ejemplo, agua. Se listarán a continuación algunas de estas propiedades:

- Los hidrocarburos transportados comúnmente tienen propiedades fisicoquímicas que dificultan su medición precisa.
- La caracterización del fluido depende directamente de su composición y su relación a la temperatura y presión que se encuentre.
- Los productos transportados son de alto valor económico, y por tanto, fuente de recursos para delincuentes.
- Por otra parte, los productos son inflamables y volátiles, lo cual su manipulación y manejo debe ser de extremo cuidado, por tanto, una fuga o pérdida no detectada puede generar emergencias graves en la zona donde este se encuentre.

Para hacer un balance de pérdida/ganancia (L/G por sus siglas en inglés) se sugiere el método clásico sugerido por API STD 2560 [6]. El análisis parte de la relación entre los volúmenes recibidos y entregados en un ducto de transporte en un tiempo determinado (e.g. días, semana, meses). Las ganancias son representadas por volúmenes que no fueron cuantificados en la entrada, mayores a las actuales recibos o menores que las actuales entregas. De hecho en sistemas totalmente controlados, no existen ganancias o pérdidas, solo pequeñas desviaciones e incoherencias de las medidas [6].

2.1.1 Ecuación de pérdida/ganancia (L/G)

La principal ecuación de L/G es representada por la siguiente ecuación donde la pérdida se representa con signo negativo para unidades de volumen o masa [6]:

$$L/G = (CI + D) - (BI + R) \quad (2.1)$$

Donde:

CI = Inventario de cierre del periodo en la línea (Closing Inventory)

BI = Inventario de inicio del periodo en la línea (Beginning Inventory)

D = Entregas (Delivery)

R = Recibos (Receipts)

Para representar el L/G en unidades porcentuales, se procede por la relación L/G por el total de entregas hechas al sistema en un tiempo establecido por 100 [6].

La información recolectada debe ser supervisada y analizada por medio de hojas de control (control charts), estas permiten analizar si la operación se encuentra en un estado normal partiendo de los límites de control preestablecidos. Estos límites son propuestos partiendo de los históricos de operación o arbitrarios, como los son los contractuales. La hoja de control debe indicar los valores de L/G porcentuales, promedio y límites ($\pm 3\sigma$ de las muestras). La información puede ser plasmada en tendencias, tablas acumulativas o tablas cruzadas para segmentos del ductos, factores de medida de los equipos de monitoreo, entre otros [6].

2.1.2 Solución de desviaciones

Uno de los retos más importantes para el análisis de L/G es la identificación de desviaciones para valores fuera de límites aritméticos preestablecidos de operación de años de los ductos. Para hacer esta inspección es necesario revisar la esta información, buscar problemas obvios, entrevistar al personal de operaciones que interactuó con la información y recorrer las instalaciones [6].

Se debe revisar cada uno de los datos que se han superado los límites de control, estos valores dan síntomas históricos de comportamientos similares. Se debe verificar que cada uno de estos datos haya sido plasmado con precisión y oportunidad en los sistemas de apoyo. Es importante conocer la naturaleza de la información que entra al sistema como la integridad de la misma información que generan los reportes finales [6]

Revisar reportes de campo, opiniones del personal y datos históricos es de suma importancia. Buscar patrones de comportamiento que están escondidos a lo largo de las

variaciones operacionales mes a mes y cruzarlos con históricos de mantenimientos o calibraciones conlleva a un análisis de impacto de problemas físicos sobre la obtención de la información. Un recorrido a las instalaciones permite verificar los diferentes procedimientos y buenas prácticas que se realizan en el día a día de operación [6].

2.1.3 Imprecisiones e incertidumbres

Muchos comportamientos y causas de la operación contribuyen imprecisiones e incertidumbres, por lo tanto, contribuyen a las pérdidas o ganancias de un sistema en particular [6]:

Medidores

- Los factores de medición (MF) son susceptibles de cambiar en función de las variaciones de la operación. Cambios en el flujo, la temperatura, la presión y la densidad implican un cambio en las condiciones en la cual fue patronado el factor como tal.
- El MF puede cambiar si el equipo de medición se encuentra en operando en condiciones no lineales.
- Un MF correcto se debe aplicar en condiciones de operación correctos. Los equipos de medición deben ser correctamente escogidos para las condiciones operacionales específicas de cada sistema.
- MF mal calculados es la causa común de una fuente de información imprecisa. Por ejemplo utilizar un MF de gasolina para medir diésel causa un error de aproximadamente 3%.
- Arranque y paradas de bombeo introducen errores, debido a que las rampas de desaceleración son más lentas que las de aceleración de unidades.
- Fugas internas o sellos defectuosos de válvulas generan volúmenes que evitan la cuantificación de los mismos por los equipos de medición.

Probadores

- Un buen probador requiere condiciones estables de operación. Sin embargo esto es posible si se hace 5 pruebas consecutivas o corridas al mismo producto, lo

cual no significa un MF real para el sistema. Por tanto se debe asegurar la repetitividad del mismo.

- Los probadores para ductos son calibrados regularmente con agua. Esta debe estar libre de aire y a una temperatura estable. Con los procedimientos propuestos por NIST se llega hasta una precisión del 0.003% al 0.004% dependiendo el tamaño y el método de calibración. Cuando se utilizan métodos de campo por canecas o serafines la incertidumbre puede llegar desde 0.05% hasta 0.15%.
- Fugas internas de las válvulas de bloqueo generan muestras incorrectas.
- Suciedad interna genera errores en la calibración del probador.

Tanques

- El certificado de medidas puede ser impreciso. El tanque puede tener un piso flexible o paredes están cubiertas de lodo o incrustaciones.
- La temperatura de los tanques puede estar incorrectamente tomada si los termómetros no son estabilizados con el equilibrio térmico del líquido contenido.
- Si el tanque es de recibo, no se drena antes de tomar las medidas de nivel.
- Las medidas son tomadas en la zona crítica de tanques con techo flotante.
- Un pozo de medición sin agujeros puede generar errores ya que se cuantificaría volúmenes muertos y temperaturas del mismo volumen.
- Los puntos de referencia de medición fueron cambiados de posición entre periodos de cuantificación.
- La precisión de las tablas de aforo de un tanque dependen directamente de la cinta con la que fue medido. La temperatura y tensión de la cinta, temperatura de lámina, condición de medidas (a fondo o vacío) y precisión de la cinta de operación.
- Volúmenes de los tanques tienden a cambiar en el tiempo. Esto se genera por factores físicos como encogimiento de las láminas por uso, deslizamiento de las láminas con los equipos de sujeción, entre otros.
- Tanques hasta de 1000bls no tienen cambios sustanciales en largo tiempo. Tanque mayores a esta capacidad deben ser re-aforados cada 10 años mínimo.

Sistema y procedimientos

- Una fuente real de pérdida es la evaporación. El espacio vacío entre el techo de un tanque y el líquido almacenado, es llenado por gases emanados por los hidrocarburos que se concentran allí. Cuando el tanque es llenado, el vapor es evacuado y por tanto perdido. Pérdidas por evaporación pueden ser reducidas por el uso de tanque con techo flotante o conexión de los venteos a sistemas de recuperación de vapor.
- Equipos de apoyo a la medición o compensación que no han sido calibrados, certificados o verificados pueden aportar al balance L/G.
- El error más común ocurre cuando se calculan tiquetes de medición manualmente por errores aritméticos o factores erróneos de corrección.
- Tiquetes que no han entrado al esquema contable causan una desviación de L/G aparentes en el periodo considerado, estos amplían u ocultan L/G.
- La lectura de cierre de un tanque de un periodo previo puede coincidir con la lectura de un tanque operativo para el mismo periodo, se traslapan las dos medidas.
- La precisión de las medidas tomadas a cierre de mes son de mucha importancia porque son utilizadas para balancear el sistema a condiciones de apertura de inventarios de mes.
- El lleno de línea contribuye significativamente al sistema de inventarios. Si es posible, el llenado de línea debe ser corregido por temperatura y presión. Las líneas deben estar completamente vacías o completamente llenas en el inicio o fin del periodo contable.
- El muestreo de las líneas o tanques debe estar uniforme para asegurar que una muestra es representativa.
- Drenajes operativos de diferentes fuentes pueden generar desviaciones importantes cuando estos son reintegrados al sistema sin una cuantificación apropiada. Por lo general estos drenajes son pequeños pero son acumulativos en un periodo largo e inciden finalmente en el inventario del sistema.
- Pérdidas aparentes pueden resultar de la reducción o mezclado de productos con diferencias significativas de densidad o químicas.

- Cambio en las características del producto, presión o temperatura de operación son indicadores de una condición de transición está sucediendo. Estas son algunas fuentes:
 - Una fuga en ducto o válvula.
 - Una inapropiada alineación del sistema.
 - Error en cálculo de volúmenes.
 - No aplicar un MF para registrar un volumen.
 - Aplicar un MF para condiciones diferentes de presión y rata de flujo.
 - Comparar una temperatura de compensación de volumen para un volumen bruto.
 - Malfuncionamiento de los medidores.
 - Malfuncionamiento de los sistemas automáticos de medida.

2.1.4 Explicación de L/G

Las desviaciones de L/G pueden ser explicadas y cuantificadas, pero en algunos casos puede ser explicada solamente. Pequeñas desviaciones sistemáticas y almacenamientos temporales pueden resultar de los siguientes eventos [6].

- Presión del ducto cambiante, incremento o decremento, esto creará una falsa condición, lo cual acumulará volumen en la línea en función de la variación de la presión.
- Interfaces de producto causan variaciones en la lectura medida por el resultado de cambio de densidades.
- Cambio de temperatura de corto y largo plazo afecta la medición por la expansión o contracción del ducto. Desbalances se producen cuando el ducto pasa por debajo de ríos y la temperatura cambia.
- Pequeñas fugas.
- Evaporación
- Encogimiento volumétrico.

Desviaciones operativas

Ejemplos de desviaciones de sistemas incluyen, pero no solamente limitados a estos [6]:

- Métodos de análisis.

- Diferentes tipos de medidores.
- Procedimientos para certificar medidores.
- Sistemas de medición, tanques vs medidores.
- Diferentes escalas de medición, Fahrenheit vs Celsius o Psia vs Psig.
- Frecuencia de pruebas de medidores.
- Propiedades de los líquidos.
- Las características físicas dada por los líquidos pueden ser presentadas por las tablas de corrección con baja precisión.
- Acumulación de parafina en las paredes del ducto cuando un crudo parafinado es llevado a su temperatura de solidificación. La parafina e impurezas introducidas por mantenimientos reduce el llenado de la línea
- Viscosidad.
- Lleno de línea.
- Errores en la capacidad de lleno de un tanque.
- Fondos de los tanques flexibles.
- Movimiento de la platina de medición.
- Una inadecuada contrapresión de medidores.

2.2 Sistemas de medición de hidrocarburos comunes en ductos

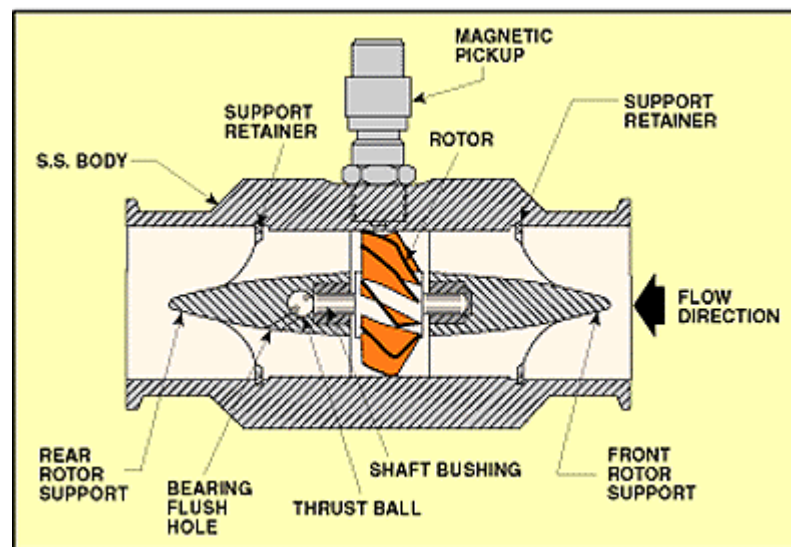
En la industria del petróleo es frecuente el uso de dispositivos para la medición de flujo para la cuantificación, monitoreo y balance en diferentes etapas de los procesos que la conforman. Para el caso particular de los ductos de transporte, los equipos más utilizados para la cuantificación son equipos de medición con turbina y para el monitoreo son equipos de medición ultrasónicos. Esta selección parte de las consideraciones operativas de los ductos tales como cambios densidad entre lotes, inspecciones en línea con equipos PIG, viscosidad variables entre otros.

En esta sección se explorará las consideraciones de cada uno de estos equipos de medición para así tener un punto de partida preciso del cómo funciona y cómo afectan estos en el uso de la industria del petróleo. Para esto nos basaremos en la normativa del API MPMS capítulo 5 secciones 3[8] y 8[9].

2.2.1 Medidores de turbina

Un medidor de turbina es un equipo de medición de flujo con un rotor que cuantifica la velocidad de un líquido que fluye en un conducto cerrado. El líquido que fluye causa que el rotor se mueva con una velocidad tangencial proporcional al promedio de la velocidad de la corriente (lo cual es cierto si la resistencia del rotor es despreciable). El promedio de la velocidad de la corriente es asumido proporcional al flujo volumétrico (lo cual es cierto si el área seccional a través del rotor se mantiene constante). El movimiento del rotor puede ser detectado mecánicamente, ópticamente o eléctricamente y es registrado. El volumen que pasa a través del medidor es determinado por pruebas contra un volumen conocido, en la industria del petróleo, se conoce como probador y se encuentran instalados en la línea de flujo con las turbinas [8].

Figura 2-1: Partes de un medidor de turbina [10].



Rendimiento del medidor de turbina

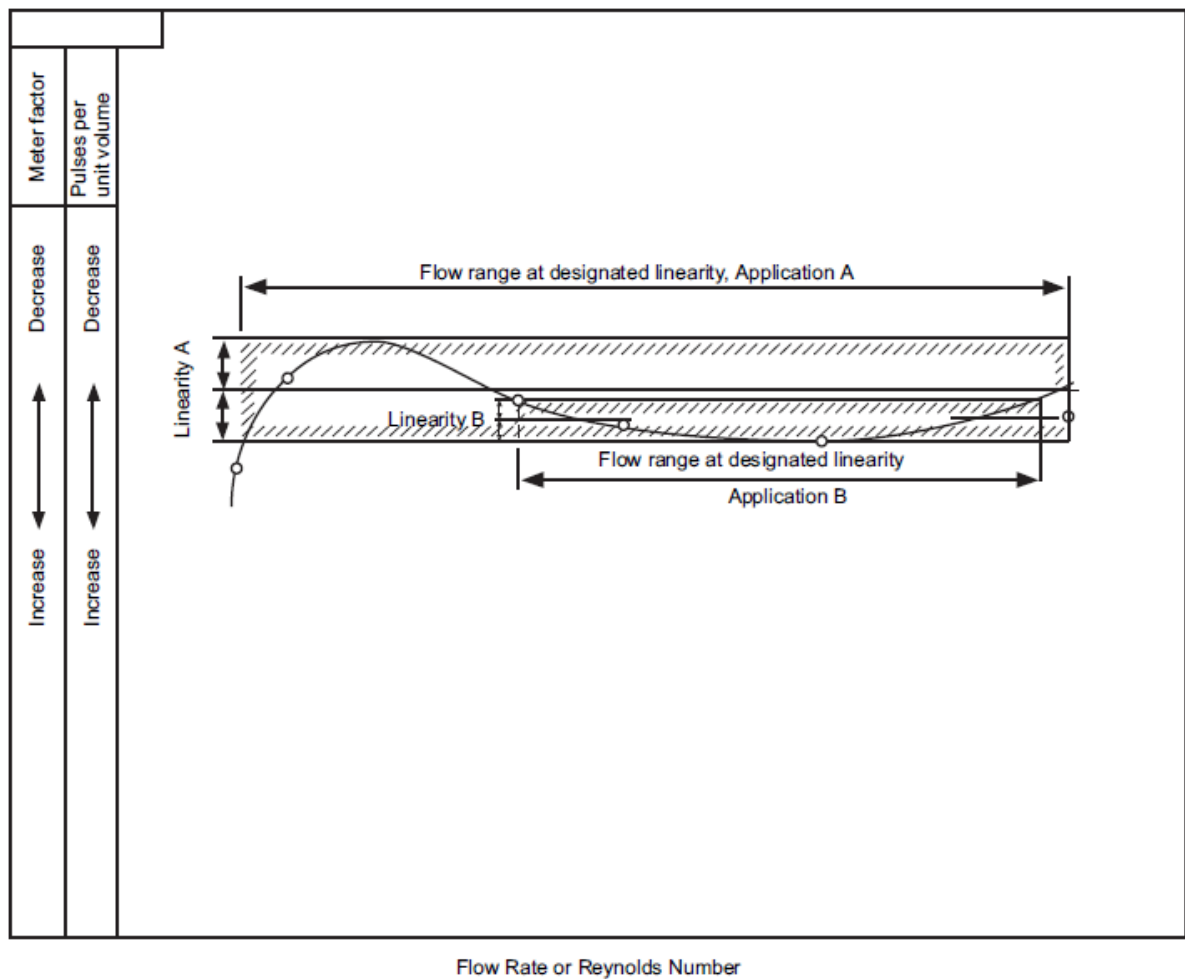
La operación del medidor está definida como que tan bien el sistema reproduce o puede hacer reproducir una cantidad medidas precisamente. Los factores de medición pueden ser determinados probando el medidor bajo condiciones de rata de flujo, viscosidad, temperatura, densidad y presión similares a las que existirán cuando se esté en operación [8].

Muchos factores pueden cambiar el rendimiento del medidor de turbina. Algunos factores, como la entrada de partículas extrañas dentro del medidor, pueden ser solo

remediados por la eliminación de la causa. Otros factores, como la deposición de partículas en el medidor, dependen de las características del líquido a ser medido; estos factores deben ser tenidos en cuenta por un diseño apropiado y la operación del sistema de medición [8].

Convencionalmente el rendimiento de turbinas multi-aspas en su rango más lineal es cuando se cuenta con un número de Reynolds (Re) sobre 30,000. Turbinas de doble aspa helicoidales operan en su rango más lineal cuando son operadas dentro de un régimen turbulento (sobre 10,000 Re). Cada medidor de turbina usualmente tiene una “curva de operación universal”, la cual es graficada del factor k o m versus el Re . Por lo tanto, si el flujo y la viscosidad son duplicados, el factor k o m del medidor en particular típicamente no tendrá un cambio significativo si el Re no cambia [8].

Figura 2-2: Rendimiento de medidores de turbina entre factores M y K [3].



Las variables que afectan a los factores del medidor (MF) son la rata de flujo, viscosidad, temperatura, depósitos y material extraño. Si un medidor es probado y operado en líquidos con propiedades idénticas, una alta precisión puede ser determinada. Si hay cambios en una o más propiedades del líquido, en las condiciones de operación y/o en las condiciones de medición internas, entre los ciclos de prueba y operación, un cambio en los factores del medidor podría resultar, obligando a determinar un MF por prueba [8].

Se listará a continuación algunos factores que impactan la medición de hidrocarburos líquidos por medidores de turbina [8].

- *Cambios en rata de flujo:* En la parte final de la curva del rango de flujo del medidor puede volverse menos linear y menos repetible que si estuviese en la parte media o alta de la misma curva. Si las gráficas de un factor de medición versus la rata de flujo ha sido desarrollado para un líquido en particular, y las

otras variables son constantes, un factor de medición puede ser seleccionado de la gráfica par flujo que se encuentren dentro de los rangos de operación del medidor; sin embargo, para mayor precisión, el medidor debe ser probado nuevamente para que opere a la nueva rata [8].

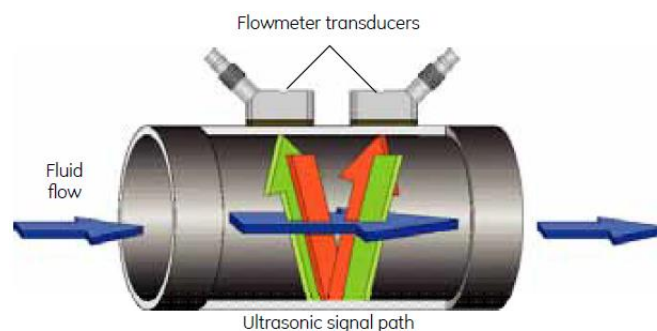
- *Cambios de viscosidad:* Los medidores de turbina son sensibles a los cambios de viscosidad. Desde que la viscosidad del hidrocarburo líquido cambie con la temperatura, la respuesta del medidor de turbina depende de ambas variables. La viscosidad en hidrocarburos livianos como la gasolina esencialmente se mantienen iguales en un espectro amplio de temperaturas. Para hidrocarburos más pesados como el petróleo crudo, el cambio en los MF pueden ser significativos porque existen cambios de la viscosidad asociados con variaciones menores de la temperatura. Es aconsejable probar el medidor frecuentemente cuando la viscosidad del fluido variará bajo las condiciones de operación normales [8].
- *Cambios de temperatura:* Adicional a los cambios de viscosidad, variaciones significativas en la temperatura de los líquidos también pueden afectar el rendimiento de la medición por cambios causados en la dimensión física del medidor. Para una mayor precisión, el medidor debe ser probado en condiciones normales de operación. Una corrección calculada de temperatura basada en el volumen promedio de temperatura de entrega, puede ser utilizado para indicar correctamente el volumen a volumen referencia o base de temperatura [8].
- *Cambio de densidad:* Para líquidos con una densidad relativa aproximadamente de 0,7 o menos, consideraciones adicionales deben ser tomadas en pro de mantener la linealidad a ratas de flujos bajos. El torque de desplazamiento de la corriente de flujo en el rotor es proporcional a la densidad del líquido multiplicada por el cuadrado de la velocidad del líquido. El torque a mínimas ratas de flujo puede ser mantenido por el incremento de la rata de flujo para flujos de baja densidad. La cantidad de incrementos en la rata de flujo mínima variará dependiendo en el tamaño y tipo de medidor. Para establecer una rata mínima de flujo, alguno pruebas deberán ser hechas a diferentes ratas hasta un factor de medición que cubra una linealidad aceptable y un repetitividad puede ser determinada [8].

- *Cambios de Presión:* Si la presión del líquido cuando es medido varía de la presión que existió durante la prueba, el volumen relativo del líquido cambiará como el resultado de su compresibilidad. Las dimensiones físicas del medidor también podrán cambiar como resultado de la expansión o contracción del chasis bajo presión. El error potencial incrementa proporcionalmente a la diferencia entre la prueba y las condiciones de operación [8].
- *Depósitos:* Depósitos en el medidor de turbina reducirán el área de flujo, lo que conlleva al incremento de la velocidad del líquido a través del rotor. Esto incrementará la velocidad, y por lo tanto el factor k del medidor, para una tasa de flujo dada. El efecto es menor en medidores de turbina helicoidales pero es substancial dependiendo del espesor de la capa y el tamaño del medidor. Depósitos en otros componentes de un medidor de turbina, o en elementos acondicionadores de flujo, también pueden ocasionar efectos significativos en el rendimiento del medidor [8].

2.2.2 Medidores ultrasónicos

Los medidores ultrasónicos son medidores inferenciales que derivan la tasa de flujo de un líquido por la medición de los tiempos de transito de pulsos de alta frecuencia. Los tiempos de transito son medidos de pulsos sonoros que atraviesan diagonalmente a través del ducto, en dirección y contra la dirección del flujo. La diferencia en esos tiempos de transito está relacionada con la velocidad promedio del flujo del líquido a lo largo de múltiples caminos acústicos. Técnicas de cálculos numéricos son utilizadas para computar el promedio de velocidad de flujo y la tasa volumétrica de flujo a las condiciones de la línea a través del medidor [9].

Figura 2-3: Esquema de montaje de transductores ultrasónicos [11].



El campo de aplicación de estos equipos y las recomendaciones dadas es en la medición dinámica de hidrocarburos líquidos, como los medidores de turbina también. Estos medidores son aptos para la medición de transferencia en custodia y otras aplicaciones como verificación de medición, chequeo de medidores y medición de detección de fugas [9].

Rendimiento de medidores ultrasónicos

Los factores de medidor deben ser determinados por una prueba del medidor en operaciones estables de medición (Rata de flujo, densidad, viscosidad, temperatura y presión constantes). Las condiciones de prueba deben ser tan cercanas a las condiciones de medición como prácticas. El propósito principal de la pruebas es confirmar el rendimiento del medidor a condiciones normales de operación. Han surgido preguntas frecuentes concernientes a las diferencias entre probar o calibrar un medidor en laboratorio o en sitio. Estos dos tipos de pruebas pueden producir diferentes resultados y no pueden necesariamente ser intercambiados sin introducir un error de medición [9].

La prueba en sitio es común y fiable dado que permite verificar el rendimiento del medidor en condiciones normales de operación. Otras condiciones de operación pueden afectar la precisión de los medidores y la repetitividad. Pruebas en sitio a condiciones estables de operación compensan variaciones en el rendimiento causado por la rata de flujo, viscosidad, densidad, temperatura, presión como también condiciones de flujo, configuración de tubería y contaminante [9].

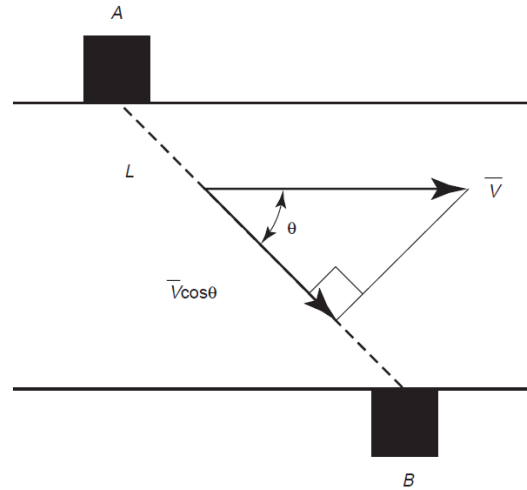
Pruebas de laboratorio no son frecuentes debido a que estas condiciones no podrían duplicar las condiciones normales de operación. Mientras existan más condiciones de incertidumbre asociadas con las pruebas de laboratorio, bajo ciertas condiciones, este puede proveer una mejor alternativa [9].

Efecto físico de medición ultrasónica

Los medidores ultrasónicos de tiempo transitorio usan transductores acústicos que envían y reciben pulsos acústicos de alta frecuencia. Estos transductores están localizados de manera que los pulsos acústicos generados atraviesen diagonalmente el ducto. Esta metodología no es sinónimo con la técnica ultrasónica Doppler que relaciona las desviaciones de mediciones de frecuencia reflejada en energía acústica [9].

La medición está basada en el hecho que el pulso acústico que atraviesa diagonalmente el ducto en dirección del flujo va a tomar menos tiempo que en atravesar que uno que atravesase en sentido opuesto. La diferencia de tiempo entre los dos pulsos acústicos es proporcional al promedio de la velocidad del flujo [9].

Figura 2-4: Modelo de medición del tiempo transitorio entre los transductores [9].



El pulso acústico que viaja en la dirección del flujo atravesar el ducto en un tiempo, ecuación 2.2 [9]:

$$t_{A \rightarrow B} = \frac{L}{c + \bar{V} \cos \theta} \quad (2.2)$$

El pulso acústico que viaja contra la dirección del flujo y atraviesa el ducto en un tiempo, ecuación 2.3 [9]:

$$t_{B \rightarrow A} = \frac{L}{c - \bar{V} \cos \theta} \quad (2.3)$$

Dónde:

L = Longitud del camino acústico.

c = Velocidad del sonido en el líquido.

θ = Ángulo del camino acústico que se hace con el ángulo del ducto.

V = Velocidad axial promedio por el ducto.

La velocidad promedio es entonces determinada por lo siguiente, ecuación 2.4 [9]:

$$\bar{V} = \frac{L}{2 \cos \theta} \times \frac{t_{B>A} - t_{A>B}}{t_{B>A} \times t_{A>B}} \quad (2.4)$$

Múltiples transductores acústicos pueden ser utilizados para crear múltiples caminos acústicos sobre la misma sección de tubería en el sentido de obtener una mejor información acerca de la distribución de la velocidad del flujo [9].

Partículas sólidas, burbujas de gas e interfaces de agua pueden generar disturbios en el pulso acústico que viaja a través del ducto. Los disturbios típicos son refracción, reflexión, atenuación y distorsión. En estos casos, la medición a la largo del camino puede ser ignorada de acuerdo al algoritmo utilizado por el fabricante. Usualmente, números bajos de mediciones ignoradas no impactarían la precisión del medidor de flujo, pero cuando existe un nivel alto de estas mediciones se puede tener un impacto considerable en la precisión del equipo, y en casos extremos detener la operación del medidor [9].

2.3 Incertidumbre

La incertidumbre es el resultado de una medición que refleja la falta de un conocimiento exacto de un valor medido. El resultado de la medición después de corrección efectos sistemáticos reconocidos es solo un valor estimado de la medición debido a que la incertidumbre proviene efectos aleatorios de la imperfecta corrección del resultado para los efectos sistemáticos [12].

En la práctica, existen posibles fuentes de incertidumbre en una medida, incluyendo [12]:

- a. Definición incompleta de la medida.
- b. Realización incompleta de la definición de la medida.
- c. Muestras no representativas; la muestra medida no representa la definición de la medida.

- d. Inadecuado conocimiento de los efectos de las condiciones ambientales de la medida o la imperfecta medición de las condiciones ambientales.
- e. Desviaciones de personal en la lectura de instrumentos análogos.
- f. Resolución finita de instrumentos o discriminación de ventana ("threshold").
- g. Valores inexactos de la medida de estándares y materiales de referencia.
- h. Valores inexactos de constantes y otros parámetros obtenidos de fuentes externas y usadas en algoritmos de reducción de datos.
- i. Aproximaciones incorporadas en el método de medida y procedimiento.
- j. Variaciones en la repetición de observaciones de la medida bajo aparentes condiciones idénticas.

Estas condiciones no son necesariamente independientes, y algunas de las fuentes desde el literal a. hasta el i. Pueden contribuir a la fuente literal j. Por su puesto que un no reconocimiento de un efecto sistemático no puede ser tomado en cuenta en la evaluación de la incertidumbre del resultado de la medición pero contribuye a este error [12].

La incertidumbre se clasifica en dos categorías basados en su método de evaluación: tipo A o tipo B. Ambas categorías se pueden aplicar el término *incertidumbre*, se debe tener en cuenta que este término no es substituto para las palabras *aleatorio* y/o *sistemático*. La incertidumbre de un efecto sistemático conocido puede en algunos casos ser obtenido por la evaluación tipo A mientras que en otros casos por la evaluación tipo B. Este también puede ser caracterizado como una incertidumbre de efectos aleatorio dado el caso [12].

En algunas publicaciones, componentes de la incertidumbre son categorizadas como aleatorias o sistemáticas. Como la categorización de los componentes de la incertidumbre pueden ser ambiguos cuando se aplica generalmente. Por ejemplo, un componente aleatorio de la incertidumbre en una medida puede volverse un componente sistemático de la incertidumbre en otra medida en la cual el resultado de la primera medida es usado como un patrón de entrada. La categorización de los métodos de la evaluación de los componentes de la incertidumbre más que los componentes en sí mismos permite ambigüedades.

El propósito de la clasificación Tipo A y B es indicar las dos diferentes maneras de evaluar los componentes de la incertidumbre y es para la conveniencia de discusión solamente; la clasificación no significa la indicación que allí exista una diferencia en la naturaleza de los componentes resultantes de los dos tipos de evaluación. Ambos tipos de evaluación están basados en la distribución de probabilidad y el resultado de la incertidumbre de los componentes del mismo tipo que son cuantificados por la variación de las varianzas de las desviaciones estándar [12].

La incertidumbre estándar Tipo A es obtenida de la función de densidad de probabilidad derivado de la distribución de frecuencia observada, mientras que incertidumbre estándar Tipo B es obtenido de una función de densidad probabilidad asumida basada en el grado de creencia que un evento ocurrirá (frecuentemente llamada probabilidad subjetiva). Ambas aproximaciones emplean interpretaciones reconocidas de probabilidad. Una evaluación de la un componente de incertidumbre es usualmente basado en la comparación de un grupo de información creíble [12].

La incertidumbre estándar es el resultado de la medición, cuando ese resultado es obtenido de los valores de un número de otras cantidades, es referido incertidumbre estándar combinada y denotada por u_c . Esa es la desviación estándar estimada asociada con el resultado y es igual a la raíz cuadrada positiva de la varianza combinada obtenida de todas las varianzas y covarianzas de los componentes [12].

Para conocer las necesidades de algunas aplicaciones industriales y comerciales, como también los requerimientos en las áreas de la salud y seguridad, una incertidumbre expandida U es obtenida por la multiplicación de la incertidumbre estándar combinada u_c por un factor de convergencia. La intención propuesta de U es la de proveer un intervalo sobre el resultado de una medida que puede ser esperada para agrupar una fracción de la distribución de los valores que pueden razonablemente atribuidos a la medida. La selección de un factor k , el cual está usualmente en el rango de 2 a 3, está basada en la convergencia de probabilidad o en el nivel de confianza requerida para el intervalo [12].

2.3.1 Consideraciones prácticas

Si el total de las cantidades en el cual el resultado de la medida depende son variadas, su incertidumbre puede ser evaluada por expresiones estadísticas. Sin embargo, porque

esto es raramente posible en la práctica dado por tiempo limitado y recursos, la incertidumbre de la medición resultante es usualmente evaluada usando un modelo matemático de la medición y la ley de la propagación de la incertidumbre. Está implícito entonces que una medida puede ser modelada matemáticamente por el grado impuesto por el requerimiento de precisión de la medición [12].

Debido a que el modelo matemático puede estar incompleto, todas las cantidades relevantes deben ser diferenciadas de la total extensión práctica llevando a que la evaluación de la incertidumbre pueda ser basada lo mejor posible en los datos observados. Siempre que sea posible, el uso de modelos empíricos de la medición partiendo de datos cuantitativos de larga duración y el uso estándares de chequeo y cartas de control que pueden indicar si una medición estaba bajo control estadístico, debe ser parte del esfuerzo para obtener evaluaciones de incertidumbre confiables. El modelo matemático debe ser siempre revisado cuando los datos revisados, incluyendo el resultado de determinaciones independientes de la misma medida, demuestren que el modelo está incompleto. Un experimento bien diseñado puede facilitar enormemente la evaluación de la incertidumbre confiable y es una parte importante del arte de la medición [12].

En el sentido de decidir si un sistema de medición está funcionando adecuadamente, la variabilidad experimentalmente observada de sus valores de salida, como mediciones por sus desviaciones estándar, es frecuentemente comparada con la desviación estándar predicha obtenida por la combinación de varios componentes de incertidumbre que caracterizan la medida. En algunos casos, solo algunos componentes (si fueron obtenidos de evaluaciones Tipo A o B) que puedan contribuir a la variabilidad observada experimentalmente de estas salidas deben ser consideradas [12].

En algunos casos, la incertidumbre por efectos sistemáticos no necesita ser incluida en la evaluación de la incertidumbre resultante. Aunque la incertidumbre haya sido evaluada, esto puede ser ignorado si su contribución para la incertidumbre estándar combinada de la medición resultante es insignificante. Si el valor de la corrección en sí mismo es relativamente insignificante para la incertidumbre estándar combinada, ésta también puede ser ignorada [12].

Esto frecuentemente ocurre en la práctica, especialmente en el dominio de la metrología legal, que un dispositivo es probado a través de una comparación con un estándar de medición y las incertidumbres asociadas con el estándar y el procedimiento de comparación son relativamente despreciables para el requerimiento de precisión de la prueba. Un ejemplo de esto es el uso de un grupo de estándares bien calibrados de masa para la prueba de precisión a escala comercial. En algunos casos, porque los componentes de incertidumbre son suficientemente pequeños para ser ignorados, la medición puede ser vista como la determinación del error del dispositivo bajo prueba [12].

La estimación del valor de una medición proveída por el resultado de una medición es a veces expresada en términos de valores adoptados de una medición de un estándar más que en términos de una unidad relevante del sistema internacional de medidas (SI). En algunos casos, la magnitud de la incertidumbre atribuible a la medición resultante puede ser significativamente pequeña que cuando los resultados son expresados en unidades relevante en unidades SI. En efecto, la medición ha sido redefinida por ser la relación entre el valor de las cantidades a ser medidas por el valor adoptado del estándar [12].

Equivocaciones en el registro y análisis de datos pueden introducir un error desconocido significativo en el resultado de la medida. Grandes equivocaciones pueden usualmente ser identificadas por una revisión apropiada de los datos; pequeños errores pueden ser enmascarados o también representados por variaciones aleatorias. Mediciones de la incertidumbre no están destinadas para tener en cuenta estos errores [12].

La evaluación de la incertidumbre no es un tarea rutinaria o un puramente matemática; esta depende del detalle de conocimiento de la naturaleza de la mediciones y de la medida. La calidad y utilidad de la incertidumbre aportada por el resultado de la medición por tanto dependerá del entendimiento, análisis crítico e integridad de quienes contribuyeron a la asignación de este valor [12].

2.3.2 Modelamiento de la medición

En la mayoría de los casos, una medición Y no es tomada directamente, pero está determinada de otras N cantidades X_1, X_2, \dots, X_N a través de la relación funcional f (ecuación 2.5) [12]:

$$Y = f(X_1, X_2, \dots, X_N) \quad (2.5)$$

Las cantidades de entrada X_1, X_2, \dots, X_N sobre la cantidad de salida Y pueden ser por si mismas ser vistas como medidas y pueden por si mismas depender de otras cantidades, incluyendo correcciones y factores de corrección por efectos sistemáticos, lo que nos lleva a una relación funcional complicada puede ser no descrita explícitamente. Más aún, f puede ser determinada experimentalmente o existir solo como un algoritmo que debe ser evaluado numéricamente [12].

Por lo tanto, si el dato indica que f no modela la medición por los grados impuestos por el requerimiento de precisión de la resultante de la medida, mediciones adicionales deben ser incluidas en f para eliminar la insuficiencia de datos. Esto requiere introducir una cantidad de entrada para reflejar el conocimiento incompleto de un fenómeno que afecta la medida [12].

Un grupo de cantidades de entrada X_1, X_2, \dots, X_N pueden ser categorizadas como [12]

- Cantidades con valores e incertidumbres directamente determinadas por una medición corriente. Estos valores e incertidumbres pueden ser obtenidas, por ejemplo, de una simple observación, observaciones repetidas o juicios basados en la experiencia. Pueden también envolver la corrección de lectura de instrumentos y correcciones por cantidades influyentes como temperatura ambiente, presión barométrica, humedad, entre otras.
- Cantidades con valores e incertidumbres incluidos por medición de fuentes externas, como cantidades asociadas con estándares de medida de calibración, certificados de referencia de material y datos de referencias obtenidas de manuales.

Un estimado de la medición Y , denotado por y , es obtenido de la anterior ecuación usando estimados de entrada x_1, x_2, \dots, x_N para los valores de N cantidades de X_1, X_2, \dots, X_N . Entonces la salida estimada y , la cual el resultado de la medición está dado por (ecuación 2.6) [12]:

$$y = f(x_1, x_2, \dots, x_N) \quad (2.6)$$

La estimación de la desviación estándar asociada con la estimación de salida o la medición resultante de y , referida como la incertidumbre estándar combinada y denotada como $u_c(y)$, está determinada de la estimación de la desviación estándar asociada con cada entrada estimada x_i , referida como incertidumbre estándar y denotada como $u(x_i)$ [12].

Cada entrada estimada x_i y su incertidumbre estándar asociada $u(x_i)$ son obtenidas de una distribución de posibles valores de la cantidad de entrada X_i . Esta distribución de probabilidad es basada en una serie de observaciones $X_{i,k}$ de X_i o puede ser una distribución *a priori*. Evaluaciones Tipo A de los componentes de incertidumbre son basados en distribuciones de frecuencia mientras que las evaluaciones Tipo B son basadas en distribuciones *a priori*. Esto debe ser reconocido que en ambos casos las distribuciones son modelos que son usados para representar el estado de nuestro conocimiento [12].

Evaluación de la Incertidumbre Estándar Tipo A

En muchos casos, el mejor estimado disponible de la expectativa o de un valor esperado μ_q de una cantidad q que varía aleatoriamente y que para cada n observaciones independientes q_k han sido obtenidos bajo las mismas condiciones de medición, es la media aritmética o promedio \bar{q} de las n observaciones (ecuación 2.7) [12].

$$\bar{q} = \frac{1}{n} \sum_{k=1}^n q_k \quad (2.7)$$

Por lo tanto, para una entrada de cantidad X_i estimada de n observaciones repetidas independientes $X_{i,k}$, la media aritmética \bar{X}_i obtenida de la ecuación 2.7 es usada como una entrada estimada x_i en la ecuación 2.6 para determinar la medición resultante y ; esto es, $x_i = \bar{X}_i$. Estos estimados de valor no evaluados de observaciones repetidas deben ser obtenidos por otros métodos, como son indicados en la segunda categoría antes dicha [12].

Las observaciones individuales q_k difieren en valor por las variaciones aleatorias en las cantidades influenciadas o efectos aleatorios. La varianza experimental de las observaciones, las cuales son estimadas de la varianza σ^2 de la probabilidad de distribución de q , esta dada por (ecuación 2.8) [12]:

$$s^2(q_k) = \frac{1}{n-1} \sum_{j=1}^n (q_j - \bar{q})^2 \quad (2.8)$$

Este estimado de varianza y su raíz cuadrada positiva $s(q_k)$, referida como desviación estándar experimental, caracterizada por la variabilidad de los valores observados q_k , o más específicamente, su dispersión su promedio \bar{q} [12].

El mejor estimado de $\sigma^2(\bar{q}) = \sigma^2/n$, la varianza de la media, esta dada por (ecuación 9) [12]:

$$s^2(\bar{q}) = s^2(q_k)/n \quad (2.9)$$

La varianza experimental de la media $s^2(q_k)$ y la desviación estándar experimental de la media $s(\bar{q})$, igual para la raíz cuadrada positiva $s^2(\bar{q})$, cuantifica que tan bien \bar{q} estima la expectativa μ_q de q y también puede ser usada como la medida de incertidumbre de \bar{q} [12].

Por lo tanto, para una entrada de cantidad X_i determinada de n observaciones repetidas independientes $X_{i,k}$, la incertidumbre estándar de su estimado $x_i = \bar{X}_i$ es $u(x_i) = s(\bar{X}_i)$ con $s^2(\bar{X}_i)$ calculado de acuerdo con la ecuación 2.9. Por conveniencia, $u^2(x_i) = s^2(\bar{X}_i)$ y $u(x_i) = s(\bar{X}_i)$ son a veces llamados Varianza Tipo A e Incertidumbre Estándar Tipo A, respectivamente [12].

El número de observaciones n debe ser lo suficientemente grandes para asegurar que \bar{q} provea una estimación confiable de la expectativa de μ_q de la variable aleatoria q y que $s^2(\bar{q})$ provean una estimación confiable de la varianza $\sigma^2(\bar{q}) = \sigma^2/n$. La diferencia entre $s^2(\bar{q})$ y $\sigma^2(\bar{q})$ debe ser considerada cuando una construye intervalos de confianza. En este caso, si la distribución de probabilidad de q es una distribución normal, la diferencia es tomada en cuenta a través de la distribución t [12].

A pesar que la varianza $s^2(\bar{q})$ es una cantidad fundamental, la desviación estándar $s(\bar{q})$ es más conveniente en la práctica porque tiene la misma dimensión que q y una mejor comprensión del valor que la de la varianza [12].

Para una medición bien caracterizada bajo control estadístico, un estimado combinado de la varianza s_p^2 que caracteriza la medición puede ser utilizada. En algunos casos,

cuando el valor de la medición q esta determinada de n -observaciones independientes, la varianza experimental del promedio aritmético \bar{q} de estas observaciones son mejor estimadas por s_p^2/n que por $s^2(q_k)/n$ y la incertidumbre estándar es $u = s_p/\sqrt{n}$ [12].

Frecuentemente un estimado x_i de una entrada de cantidad X_i es obtenido de la curva que ha sido descrita por datos experimentales por el método de mínimos cuadrados. Las varianzas estimadas y el resultado de las incertidumbres estándar de los parámetros adquiridos caracterizando la curva y de cualquier punto de predicción pueden usualmente ser calculados por procedimientos estadísticos bien conocidos [12].

Los grados de libertad v_i de $u(x_i)$, igual a $n - 1$ en el caso simple donde $x_i = \bar{X}_i$ y $u(x_i) = s(\bar{X}_i)$ son calculados de n observaciones independientes, siempre deben ser tomadas cuando las evaluaciones de componentes de incertidumbre Tipo A están documentadas [12].

Si las variaciones aleatorias en las observaciones de un entrada de cantidad están correlacionadas, por ejemplo en tiempo, la media y la desviación estándar experimental pueden ser inapropiados estimadores de las estadísticas deseadas. En algunos casos, las observaciones deben ser analizadas por métodos estadísticos especialmente diseñados para tratar series correlacionadas, medidas aleatoriamente variantes [12].

La discusión de la evaluación de la incertidumbre estándar Tipo A no significa ser exhaustiva; hay muchas situaciones, algunas más complejas, que pueden ser tratada por métodos estadísticos. Un importante ejemplo es el uso de diseños de calibración, frecuentemente basados en métodos de mínimos cuadrados, para evaluar incertidumbres provenientes de variaciones de cortas o largas duraciones en el resultado de las comparaciones de materiales de artefactos de valores desconocidos, bloques de medida y estándares de masa, con estándares de referencia de valores conocidos [12].

Evaluación de la Incertidumbre Estándar Tipo B

Para un estimado x_i de una entrada de cantidad X_i que no haya sido obtenido de observaciones repetidas, la varianza asociada estimada $u^2(x_i)$ o la incertidumbre estándar $u(x_i)$ esta evaluado por juicios científicos basados en toda la información disponible en la variabilidad de X_i . El conjunto de la información puede incluir [12]:

- Datos de medidas previas.
- Experiencia o conocimiento general del comportamiento o propiedades de materiales relevantes e instrumentos.
- Especificaciones del fabricante.
- Datos provenientes de calibraciones y otros certificados.
- Incertidumbre asignada por datos de referencia tomados de manuales.

Por conveniencia $u^2(x_i)$ y $u(x_i)$ evaluadas en este sentido son a veces llamadas Varianza Tipo B e Incertidumbre estándar Tipo B [12].

Cuando x_i es obtenido de una distribución a priori, la varianza asociada está apropiadamente escrita como $u^2(X_i)$ pero por simplicidad se utiliza la notación $u^2(x_i)$ y $u(x_i)$ [12].

El apropiado uso del conjunto de información disponible para la evaluación de la incertidumbre estándar Tipo B requiere de una visión basada en la experiencia o el conocimiento general; es una habilidad que se puede aprender con la práctica. Esta debe ser reconocida que para una evaluación de la Incertidumbre Estándar Tipo B puede ser tan confiable como una de Tipo A, especialmente en situaciones de medición donde la evaluación Tipo A esta basada en la comparación de un número pequeño de observaciones estadísticamente independientes [12].

Si el estimado x_i es tomado de las especificaciones de fabricante, certificado de calibración, manual de funcionamiento u otra fuente y su incertidumbre asignada indica se múltiplo de la desviación estándar, la incertidumbre estándar $u(x_i)$ es simplemente el valor asignado dividido por el múltiplo y la varianza estimada $u^2(x_i)$ es la raíz cuadrada de esa asignación [12].

La incertidumbre asignada de x_i no es necesariamente tomada como un múltiplo de la desviación estándar como se dijo en el párrafo anterior. De otra manera, uno puede encontrar este valor de la incertidumbre asignada definido en un intervalo del 90%, 95% o 99% del nivel de confianza. A menos que de otra manera indique, se puede asumir que una distribución normal fue usada para calcular la incertidumbre asignada y recuperar la incertidumbre estándar de x_i por la división de la incertidumbre asignada por un factor

apropiado para la distribución normal. Los factores correspondientes a estos tres niveles de confianza son 1,64; 1,96 y 2,58 [12].

Considerando el caso donde, basados en la información disponible, uno puede establecer que “allí está un cincuenta-cincuenta oportunidad que el valor de la entrada de cantidad X_i varía en un intervalo a_- hasta a_+ ” (En otras palabras, la probabilidad que X_i este entre este intervalo es del 0,5 al 50 por ciento). Si puede ser asumido que la distribución de posibles valores de X_i es aproximadamente normal, entonces el mejor estimado x_i de X_i puede ser tomado por el punto medio del intervalo. Más aún, si la media-banda del intervalo esta denotada como $a = (a_+ - a_-)/2$, se puede tomar $u(x_i) = 1,48a$ porque la distribución normal con expectativa μ y la desviación estándar σ del intervalo $\mu \pm \sigma/1,48$ abarca aproximadamente el 50 por ciento de la distribución [12].

Considerando un caso similar al anterior pero donde basados en la información disponible, se puede establecer que “existe la posibilidad de dos de las terceras partes que el valor de X_i varía en el intervalo a_- hasta a_+ ” (En otras palabras, la probabilidad de que X_i varíe dentro de este intervalo es de 0,67). Se puede entonces razonablemente tomar $u(x_i) = a$ porque para una distribución normal con expectativa μ y la desviación estándar σ del intervalo $\mu \pm \sigma$ abarca aproximadamente el 68,3 por ciento de la distribución [12].

En otros casos, puede ser posible solo estimar los límites para una X_i en particular, establecer que “la probabilidad que el valor de X_i varíe entre un intervalo a_- hasta a_+ para todos los propósitos prácticos es igual a uno y que la probabilidad que X_i varíe por fuera de este intervalo es esencialmente cero”. Si no es un conocimiento específico sobre los posibles valores de X_i dentro del intervalo, se puede solo asumir que es igualmente probable para X_i variar en cualquier lugar dentro del intervalo (una distribución rectangular o uniforme). Por lo tanto x_i , la expectativa o el valor esperado de X_i es el punto intermedio del intervalo, $x_i = (a_+ + a_-)/2$ con su varianza asociada (ecuación 2.10) [6]:

$$u^2(x_i) = (a_+ - a_-)^2/12 \quad (2.10)$$

Si la diferencia entre los límites, $a_+ - a_-$ está denotada por $2a$, entonces la ecuación 2.10 se convierte en (ecuación 2.11) [12]:

$$u^2(x_i) = a^2/3 \quad (2.11)$$

Límites superiores e inferiores a_+ y a_- para la entrada de cantidad X_i puede ser no simétrico con respecto al mejor estimado u_i ; mas específicamente si el límite inferior está escrito como $a_- = x_i - b_-$ y el límite superior como $a_+ = x_i + b_+$, siendo $b_- \neq b_+$. En este caso, x_i no es el valor central del intervalo a_- hasta a_+ y la distribución de probabilidad X_i no puede ser uniforme a través del intervalo. Sin embargo, esto puede no ser suficiente información disponible para seleccionar una distribución apropiada; diferentes modelos guiarían diferentes expresiones para la varianza. En ausencia de dicha información, la aproximación más simple es (ecuación 2.12) [12]:

$$u^2(x_i) = \frac{(b_+ - b_-)^2}{12} = \frac{(a_+ - a_-)^2}{12} \quad (2.12)$$

Es importante no tener una doble contabilización de componentes de incertidumbre. Si un componente de incertidumbre proveniente de un efecto particular y es obtenido de una evaluación Tipo B, este debe ser incluido como un componente independiente de incertidumbre en el cálculo de la desviación estándar combinada para la medición resultante únicamente para que el efecto no contribuya a la variabilidad observada de las observaciones. Esto es debido a que la incertidumbre parte del efecto que contribuye a la variabilidad observada ya se encuentra incluido en el componente de incertidumbre obtenida a partir del análisis estadístico de las observaciones [12].

2.4 Métodos para la fusión de la información

La fusión de información es un campo relativamente nuevo con un número de definiciones incompletas y más aún en el campo del transporte de hidrocarburos. Estas definiciones que están siendo utilizadas en campos como la defensa y energía que podrían ser utilizadas en la industria petrolera dado que los esquemas de transporte y logística de dichas industrias son similares para el transporte de hidrocarburos. Se usará la fusión de datos para combinar los datos producidos por uno o más sensores de manera que se pueda tener una mejor estimación de la cantidad que se está midiendo en el transporte. Las ideas actuales de la fusión de datos son dominados por dos aproximaciones: Teoría de Bayes y Teoría de Dempster-Shafer. La teoría de Bayes está basada en las ideas clásicas de probabilidad mientras que la Dempster-Shafer es una

concepción nueva que permite una mejor interpretación de la incertidumbre como un todo.

2.4.1 Filtro Kalman

Este filtro es por mucho el más familiar y algoritmo ampliamente utilizado en aplicaciones de fusión de información. Se asume que dentro de un momento k , un sensor recolecta una secuencia de tiempo $Z^k: z_1, \dots, z_k$ de vectores de observación. En cada paso de tiempo k y la base de los datos disponibles Z^k , se quiere determinar dos cosas [13]:

- Un estimado del vector de estado $x_{k|k}$.
- Un estimado del error probable asociado con $x_{k|k}$.

El vector de estado $x_{k|k}$ contiene la información sobre muestra que deseamos saber. La medición del error in $x_{k|k}$ es la matriz de covarianza $P_{k|k}$, y mide el error en ese sentido. Para cada entero $k > 0$ (ecuación 2.13) [13],

$$(x - x_{k|k})^T P_{k|k}^{-1} (x - x_{k|k}) \leq 1 \quad (2.13)$$

Esta ecuación define la híper-elipse de todo los estados x que están localizados dentro de un sigma de $x_{k|k}$ y es por lo tanto una generalización multidimensional del concepto de una barra de error [13].

El filtro Kalman nos muestra la propagación recursiva de $x_{k|k}$ y $P_{k|k}$ a través del tiempo, lo que resulta en la secuencia de tiempo (ecuación 2.14) [13],

$$(x_{0|0}, P_{0|0}) \rightarrow (x_{1|0}, P_{1|0}) \rightarrow (x_{1|1}, P_{1|1}) \rightarrow \dots \rightarrow (x_{k+1|k}, P_{k+1|k}) (x_{k+1|k+1}, P_{k+1|k+1}) \quad (2.26)$$

Iniciación del filtro Kalman

Se debe iniciar el filtro escogiendo un valor supuesto de $x_{0|0}$ sobre un valor de estado de la muestra y un valor supuesto de $P_{0|0}$ sobre el grado de certeza en ese estado. Si $x_{0|0}$ y $P_{0|0}$ son pobremente seleccionados, el filtro Kalman no será capaz de adquirir la muestra y su divergencia [13].

Predicción del filtro Kalman

Al momento recursivo de $k + 1$ se recolectará una nueva observación Z_{k+1} . Desde que la muestra se haya movido durante el intervalo entre los tiempos k y $k + 1$, se debe generar una incertidumbre creada por este movimiento. Esto es posible utilizando el *modelo formal movimiento estadístico* de la forma (ecuación 2.15) [13],

$$X_{k+1} = F_k x + V_k \quad (2.15)$$

Este modelo formal de movimiento consiste en dos partes. La primera es un modelo de movimiento determinístico, $X_{k+1} = F_k x$, el cual esencialmente es un supuesto de como la muestra se moverá entre los intervalos de tiempo k y $k + 1$. Por ejemplo, se puede suponer que la muestra se moverá en línea recta con velocidad constante o la muestra se moverá como un objeto balístico. El modelo determinístico de estados que la muestra tendrá en el estado x_{k+1} en el paso de tiempo $k + 1$ si este tiene un estado x en el paso de tiempo k . La matriz F_k es llamada la *matriz de transición de estados* [13].

Desde que el movimiento es desconocido y $X_{k+1} = F_k x$ es solo un supuesto, el actual x_{k+1} no será usualmente $F_k x$ pero, más aun, podrá ser alguna perturbación $F_k x + \Delta X$ de él. Se asume que esta perturbación, el ruido de planta, es un vector gaussiano aleatorio con promedio cero $\Delta X = V_k$ con covarianza Q_k . Si el movimiento actual de la muestra se desvía significativamente del modelo determinístico, usualmente se puede cubrir esta desorientación haciendo V_k más ruidosa. El hecho F_k, V_k puede ser escogido de una manera diferente como la muestra maniobra y contabilizada escogiendo un F_k aplicable [13].

Dado esto, la ecuación de predicción del filtro Kalman nos dice que un tiempo $k + 1$, el estado de la muestra más probable está dado por (ecuación 2.16) [13],

$$x_{k+1|k} = F_k x_{k|k} \quad (2.16)$$

Y que el posible error en esta predicción es medida por la matriz de predicción de error de covarianzas (ecuación 2.17) [13],

$$P_{k+1|k} = F_k P_{k|k} F_k^T + Q_k \quad (2.17)$$

Corrección del filtro Kalman

Teniendo una extrapolación del estado de la muestra en el tiempo de la siguiente muestra en el tiempo $k + 1$, se necesita la corrección de esta predicción usando la

observación actual z_{k+1} . Para hacer esto, se debe tener un modelo de como el sensor genera las observaciones. Esto se logra utilizando un modelo de medición de la forma (ecuación 2.18) [13],

$$Z_{k+1} = H_{k+1}x + W_{k+1} \quad (2.18)$$

Como el modelo de movimiento, este consiste de dos partes. La primera es un *modelo transformación estado a medición* determinístico $z = H_{k+1}x$, este estado que si una muestra tiene un estado x en el tiempo $k + 1$, entonces este generará observaciones z . La matriz $M \times N$ de H_{k+1} captura el hecho que el sensor no puede observar las cantidades x estados pero, más aún, solo alguna visión transformada e incompleta de la muestras. Adicionalmente, debido al ruido interno el sensor no colectará $H_{k+1}x$ pero de otra manera alguna perturbación aleatoria $H_{k+1}x + \Delta Z$. Se asume que $\Delta Z = W_{k+1}$ es un vector gaussiano aleatorio con promedio cero con matriz de covarianza R_{k+1} [13].

Dado esto las ecuaciones de corrección del filtro Kalman dicen que el mejor estimado del estado de la muestra es (ecuación 2.19) [13],

$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1}(z_{k+1} - x_{k+1|k}) \quad (2.19)$$

Y que el posible error en este estimado es medido por la matriz de error de covarianzas corregidas (ecuación 2.20) [13],

$$P_{k+1|k+1} = (I - K_{k+1}H_{k+1})P_{k+1|k} \quad (2.20)$$

Donde, la matriz $N \times M$ de ganancia Kalman K_{k+1} es (ecuación 2.21) [13],

$$K_{k+1} = P_{k+1|k}H_{k+1}^T(H_{k+1}P_{k+1|k}H_{k+1}^T + R_{k+1})^{-1} \quad (2.21)$$

La *innovación* $s_{k+1} = z_{k+1} - H_{k+1}x_{k+1|k}$ indica el grado en el cual la medida actual z_{k+1} difiere de la predicción de medición $H_{k+1}x_{k+1|k}$, y K_{k+1} determina el grado en el cual la predicción de estado $x_{k+1|k}$ debe ser corregido para reflejar esta desviación [13].

2.4.2 Fusión de la información usando filtro Kalman (KL)

Desde un modelo diferente de H_{k+1} , W_{k+1} puede ser escogido en cada paso de tiempo, el filtro Kalman provee una manera de fusionar datos provenientes de diferentes sensores. Por ejemplo, si hay observaciones estadísticamente independientes \dot{z}_{k+1} y \ddot{z}_{k+1} de la muestra y recolectadas de dos sensores al mismo tiempo, una puede ser el primer proceso \dot{z}_{k+1} usando la ecuación de corrección para el modelo $\dot{Z}_{k+1} = \dot{H}_{k+1}x + \dot{W}_{k+1}$; y entonces el proceso \ddot{z}_{k+1} usando la ecuación de corrección para el modelo $\ddot{Z}_{k+1} = \ddot{H}_{k+1}x + \ddot{W}_{k+1}$ [13].

Si los sensores no son estadísticamente independientes, entonces la fusión de sus datos requieren la construcción de un modelo de medición conjunta (ecuación 2.22) [13].

$$\begin{pmatrix} \dot{z}_{k+1} \\ \ddot{z}_{k+1} \end{pmatrix} = \begin{pmatrix} \dot{H}_{k+1} \\ \ddot{H}_{k+1} \end{pmatrix} x + \begin{pmatrix} \dot{W}_{k+1} \\ \ddot{W}_{k+1} \end{pmatrix} \quad (2.22)$$

Para las observaciones conjuntas $(z^T, \ddot{z}^T)^T$, donde la matriz de las correlaciones conjuntas de \dot{W}_{k+1} y \ddot{W}_{k+1} es tomado para el modelo la dependencia estadística de los dos sensores. Esto incrementa el tamaño del espacio de las medidas, con la consecuencia que la carga computacional para la inversión de la matriz de ganancia Kalman puede incrementarse sustancialmente [13].

2.4.3 Fusión Bayes para datos simples (BY)

Suponiendo que dos sensores estadísticamente independientes con sus respectivas funciones de probabilidad $\dot{f}_{k+1}(\dot{z}|x)$ y $\ddot{f}_{k+1}(\ddot{z}|x)$ realizó observaciones \dot{z}_{k+1} y \ddot{z}_{k+1} , respectivamente. Estas mediciones pueden ser fusionadas por la aplicación ecuación de corrección de Bayes secuencialmente, primero para \dot{z}_{k+1} y luego para \ddot{z}_{k+1} (ecuaciones 2.23-2.25) [8]:

$$f_{k+1|k+1}(x|Z^k, \dot{z}_{k+1}) \propto \dot{f}_{k+1}(\dot{z}_{k+1}|x) \cdot f_{k+1|k}(x|Z^k) \quad (2.23)$$

$$f_{k+1|k+1}(x|Z^k, \dot{z}_{k+1}, \ddot{z}_{k+1}) \propto \ddot{f}_{k+1}(\ddot{z}_{k+1}|x) \cdot f_{k+1|k+1}(x|Z^k, \dot{z}_{k+1}) \quad (2.24)$$

$$f_{k+1|k+1}(x|Z^k, \dot{z}_{k+1}, \ddot{z}_{k+1}) \propto \dot{f}_{k+1}(\dot{z}_{k+1}|x) \cdot \ddot{f}_{k+1}(\ddot{z}_{k+1}|x) \cdot f_{k+1|k}(x|Z^k) \quad (2.25)$$

Esto es equivalente a aplicar la ecuación de corrección en un tiempo simple, usando la función de probabilidad conjunta (ecuación 2.26) [8]:

$$\ddot{f}_{k+1}(\dot{z}_{k+1}, \ddot{z}_{k+1}|x) \triangleq \dot{f}_{k+1}(\dot{z}_{k+1}|x) \cdot \ddot{f}_{k+1}(\ddot{z}_{k+1}|x) \quad (2.26)$$

Si los sensores no son independientes entonces las mediciones pueden ser fusionadas entregando la función de probabilidad conjunta para el conjunto de sensores [13].

Fusión de múltiples medidores por teoría de Bayes

Siguiendo el mismo proceso propuesto en la sección 2.4.2, se debe realizar los mismos cálculos de intersección de valores contra el valor inicial y final. Se fusionarían cuantos medidores sean necesarios bajo la siguiente ecuación que aplica para tres medidores y puede ser expandida de igual manera para los medidores que deben ser fusionados (ecuación 2.27) [7].

$$(x|Y_1^1 Y_1^2 Y_1^3) = \frac{(x|Y_1^1)(x|Y_1^2)(x|Y_1^3)(x|Y_0^1 Y_0^2 Y_0^3)}{(x|Y_0^1)(x|Y_0^2)(x|Y_0^3)} \quad (2.27)$$

2.4.4 Fusión Dempster-Shafer para datos simples (DS)

Las aproximaciones de Bayes y Dempster-Shafer son basadas en el concepto asignar pesos a estados definidos del sistema que va a ser medido. Mientras que Bayes aplica un significado más clásico de ideas conocidas de la probabilidad, DS permite otros escenarios alternativos para el sistema, como tratar igualmente los conjuntos de alternativas que tiene una intersección que no pasa por cero. Pero estos pesos, lo cual en la teoría de Bayes clásica son probabilidades, son menos entendidos en la teoría de DS. Las cantidades análogas de DS son llamadas masas y pueden entenderse como probabilidades también [7].

La teoría DS asigna estas masas para todos los subgrupos de entidades que conforman el sistema. Suponga que se tiene un sistema con cinco miembros. Podemos identificarlos a todos ellos y describir como un subgrupo particular como “1” para cada elemento que está en el subgrupo. De esta manera se observa que allí hay un número 2^5 de subgrupos posibles. Si es grupo original es llamado S entonces el grupo de todos los subgrupos es llamado 2^S , grupo focal [7].

Una asignación básica de las masas (b.m.a.) en un grupo finito \mathfrak{U} es una función no negativa $m(U)$ definida en el subgrupo $U \subseteq \mathfrak{U}$ donde: 1. $m(U) = 0$ para todo número finito de U (llamado el grupo focal de m); 2. $\sum_U m(U) = 1$ y 3. $m(\emptyset) = 0$. La cantidad $m(U)$ es interpretada como el grado de credibilidad en la hipótesis U que no se acumula a más constantes de la hipótesis $V \subseteq U$. La selección $U = \mathfrak{U}$ indica una incertidumbre completa y es llamado “hipótesis nula” [13].

Expresado en una notación diferente, un b.m.a. consiste en un número finito de grupos focales U_1, \dots, U_d de \mathfrak{U} , a las cuales se le han asignado masas no negativas m_1, \dots, m_d y con $m_1 + \dots + m_d = 1$. Por lo tanto, $m(U) = 0$ a menos que $U = U_i$ para algún i , en cual caso $m(U_i) = m_i$ [13].

De un b.m.a m se pueden construir las funciones de credibilidad, la plausibilidad y homogeneidad (commonality) respectivamente (ecuaciones 2.28, 2.29 y 2.30) [13].

$$Bel_m(U) \triangleq \sum_{V \subseteq U} m(V) \quad (2.28)$$

$$Pl_m(U) \triangleq 1 - Bel_m(U^c) = \sum_{V \cap U \neq \emptyset} m(V) \quad (2.29)$$

$$Q_m(U) \triangleq \sum_{V \supseteq U} m(V) \quad (2.30)$$

Un b.m.a puede ser recuperada una m de $Bel_m(U)$ vía la transformada de Möbius (ecuación 2.31) [13].

$$m(U) = \sum_{V \subseteq U} (-1)^{|U-V|} \cdot Bel_m(V) \quad (2.31)$$

Combinación o fusión de Dempster-Shafer

Asumiendo que hay independencia en algún sentido, b.m.a. pueden ser fusionadas usando la combinación de Dempster-Shafer (ecuación 2.32) [13]:

$$(m_1 * m_2)(U) \triangleq \alpha^{-1} \sum_{U_1 \cap U_2 = U} m_1(U_1) \cdot m_2(U_2) \quad (2.32)$$

Si $U \neq \emptyset$ y $(m_1 * m_2)(\emptyset) \triangleq 0$ de otra manera. La combinación $m_1 * m_2$ está definida solo si su argumento no es cero (ecuación 33) [13]:

$$0 \neq \alpha \triangleq \alpha_{DS}(m_1, m_2) \triangleq \sum_{U_1 \cap U_2 \neq \emptyset} m_1(U_1) \cdot m_2(U_2) \quad (2.33)$$

Si solo los subgrupos focales de m_2 tiene instancia única, entonces lo mismo es cierto de $m_1 * m_2$ y tenemos la relación en la cual se parece a la regla de Bayes (ecuación 2.34) [13]:

$$(m_1 * m_2)(\{u\}) \propto L_{m_1} \cdot m_2(\{u\}) \quad (2.34)$$

Donde $L_{m_1}(u) \triangleq \sum_{U \ni u} m_1(U)$. Si adicionalmente solo los subgrupos focales de m_1 tiene instancia única entonces (ecuación 2.35) [13]:

$$(m_1 * m_2)(u) \propto m_1(u) \cdot m_2(u) \quad (2.35)$$

Fusión de múltiples medidores por Dempster-Shafer.

Siguiendo el mismo proceso propuesto en la sección 2.4.1, se debe realizar los mismos cálculos de intersección de valores contra el valor inicial y final. Se fusionarían cuantos medidores sean necesarios bajo la siguiente ecuación que aplica para tres medidores y puede ser expandida de igual manera para los medidores que deben ser fusionados (ecuación 2.36) [7].

$$m^{1,2,3}(D) = \frac{\sum_{A \cap B \cap C = D} m^1(A)m^2(B)m^3(C)}{\sum_{A \cap B \cap C \neq \emptyset} m^1(A)m^2(B)m^3(C)} = \frac{\sum_{A \cap B \cap C = D} m^1(A)m^2(B)m^3(C)}{1 - \sum_{A \cap B \cap C = \emptyset} m^1(A)m^2(B)m^3(C)} \quad (2.36)$$

2.4.5 Fusión por promedio ponderado de confianza

Para combinar medidas, una forma simple es promediar todos los valores que usualmente no es la mejor forma de hacerlo. La razón radica es que algunos sensores pueden ser más confiables que otros, así que sus medidas pueden ser representadas más precisamente. Esta es la razón por la cual se le asigna una mayor importancia y mayor peso ponderado a una observación x_i de un sensor que es más confiable que uno de menos precisión y se calcula su peso ponderado promedio como (ecuación 2.37) [14]:

$$x_{fused} = \sum_{i=1}^N x_i w_i \quad (2.37)$$

Una vía para determinar los pesos de w_i para cada uno de promedios ponderados es basado en la varianza de los errores medición hecho por cada sensor en la ecuación 2.37. La varianza de la fusión resultante puede ser determinada a través de la ecuación 2.38. Para el caso que los errores sean independientes, asignar los pesos en este sentido es óptimo dado que la varianza esperada del resultado fusionado es minimizada [14].

$$w_i = \frac{1}{\sigma_i^2 \sum_{j=1}^n \frac{1}{\sigma_j^2}} \quad (2.38)$$

$$\sigma_{fused}^2 = \sum_{i=1}^n w_i^2 \sigma_i^2 \quad (2.39)$$

La suposición de independencia de los errores de los sensores no debe ser hecha en casos generales. Ignorar correlaciones entre entradas fusionadas resulta en aseveraciones poco adecuadas de los pesos de cada observación y distorsiona el estimado de la varianza resultante. En el caso que el sistema use fuentes distribuidas de información y se fusionen datos en varios puntos del proceso, el estimado de la varianza fusionada de los valores fusionados determinara su influencia en procesos posteriores de fusión. Correlaciones positivas, por ejemplo, causa que la varianza calculada de acuerdo con la ecuación 2.39 para subestimar la varianza real, lo cual resulta en una estimación de mayor peso en el próximo proceso de fusión que debe ser asignado [14].

2.5 Definición de PDF para mediciones

Una definición apropiada de las diferentes entradas a los sistemas es necesaria para un análisis estadístico y funciona con las necesidades que presentan la identificación de pérdidas en las líneas de transporte de hidrocarburos. Las definiciones parten de las funciones de densidad de probabilidad (PDF por su acrónimo en inglés) cada uno de los equipos generadores de información que conforman el sistema a analizar. Los PDF de cada instrumento dependen directamente del instrumento en sí, su instalación y la variable física a medir, lo que hace que generalizar un PDF por grupo de instrumentos no es viable. Por tanto, definir el PDF por cada instrumento es el punto inicial para el análisis de las pérdidas de hidrocarburos de un sistema.

Dado que las diferentes entradas de información al sistema ya se encuentran instaladas en campo y la alta variabilidad de las composiciones de los elementos a medir, se tomará el método de Monte Carlo para aproximar PDF expuesto en la guía: “Guide to the expression of uncertainty in Measurement-Supplement 2” de la Organización Internacional de Metrología Legal- OIML en su capítulo 7 [15].

2.5.1 Método Monte Carlo (MCM)

El MCM provee una aproximación general para obtener una representación numérica aproximada de G , es decir, de la función de distribución $G_Y(\theta)$ para Y . El corazón de la aproximación es hacer repetidas gráficas de los PDF de X_i (o PDF conjunto de X) y la evaluación cuantitativa de un vector de salida en cada caso [15].

Desde que $G_Y(\eta)$ abarque toda la información conocida sobre Y , cualquier propiedad de Y como expectativa, varianza y covarianza, y regiones de cobertura pueden ser aproximadas utilizando G . La calidad de estos resultados calculados generalmente mejora con el número de gráficos hechos [15].

Las cantidades del grupo de salidas obtenidas en posterior al análisis de los PDF de X_i serán tomados independientemente para generar la distribución de probabilidad conjunta de Y . Expectativas, variancias y covariancias pueden ser determinados directamente de estos valores. La determinación de las regiones de cobertura requiere estos valores para ser analizados dentro de una manera apropiada [15].

Tomando y_r para $r = 1, \dots, M$ representan los valores de salida de PDF de X_i . La expectativa $E(Y)$ y varianza $V(Y)$ de Y pueden ser aproximadas usando y_r . En general, los momentos de Y (incluyendo $E(Y)$ y $V(Y)$) son aproximados por los momentos de estos valores. Tomando M_{y_0} para denotar el número de y_r para los cuales cada componente no es mayor a su componente correspondiente y_0 , descritos dentro de un vector numérico de dimensión $m \times 1$. La probabilidad $\Pr(Y \leq y_0)$ es aproximada por M_{y_0}/M . En ese sentido y_1, \dots, y_M provee una representación discreta de la función de distribución $G_Y(\eta)$ [15].

G , la principal salida de MCM, constituye la matriz de dimensión $m \times M$ dada por (ecuación 2.40) [15]:

$$G = (y_1, \dots, y_M) \quad (2.40)$$

Es solo posible la implementación de un MCM para definir la propagación de las distribuciones Y cuando puede ser expresada en términos de X donde M es conocida con antelación. Para estos casos un MCM puede ser establecido como un procedimiento paso a paso según lo siguiente [15]:

- a) Seleccionar un número M de ensayos de Monte Carlo para ser analizados.
- b) Generar M vectores mediante diagramación aleatoria de los PDF asignados a las entradas de cantidad X_i (o de PDF conjunta de X), como datos realizados para X_i .
- c) Para cada uno de estos vectores, forme valor correspondiente de Y , llenando el vector M con las cantidades dentro.
- d) Tome la representación de G de la función de distribución para Y como el grupo del vector de cantidades de salida M .
- e) Use G para formar un estimado de y de Y y la covarianza de la matriz U_y asociada con y .
- f) Use G para formar una región de cobertura apropiada para Y , para estipular una probabilidad de cobertura p .

La efectividad de MCM para determinar y , U_y y una región de cobertura Y depende del uso de un adecuado valor de ensayos M [15].

Número de ensayos de Monte Carlo

Un valor de M , el número de pruebas de Monte Carlo, debe ser seleccionado. Este puede ser escogido *a priori*, en cuyo caso no habrá un control directo sobre la calidad de los resultados numéricos proporcionados por MCM. La razón es que el número de ensayos necesarios para proporcionar estos resultados a una tolerancia numérica prescrita dependerá de la forma de la PDF para la cantidad de salida y la probabilidad de cobertura requerida. También, los cálculos son estocásticos naturalmente, siendo la base de los diagramas aleatorios [15].

Dado que un número pre-asignado de ensayos no garantiza la suficiencia de la prueba, se puede seleccionar un M adaptativamente. Se debe tener en cuenta factores

económicos en el uso de los recursos en función de las expectativas de la calidad de los resultados esperados [15].

Hacer diagramas de las distribuciones de probabilidad

En una implementación de MCM, M vectores $x_r = 1, \dots, M$ son diagramados de los PDF $g_{X_i}(\epsilon_i)$ para las cantidades de entrada X_1, \dots, X_N . Los diagramas podrán ser hechos de PDF conjuntas (multi-variables) $g_X(\epsilon)$ si es apropiado [15].

Recomendaciones concernientes a la manera con la cual estos gráficos pueden ser obtenidos por distribuciones comunes, incluyendo rectangular, Gaussiana, distribución-t y Gaussiana multivariada pueden encontrarse dentro de la misma norma en la sección 5.3.2.4. [15].

Evaluación de la cantidad del vector de salida

La cantidad del vector de salida es evaluada para cada gráfica M de los PDFs para las N cantidades de entrada. Específicamente, se denota gráficas M por $x_1 \dots x_M$, para cada dimensión $N \times 1$, donde la r -ésima gráfica x_r contiene $x_{1,r} \dots x_{N,r}$ con $x_{i,r}$ una gráfica de la PDF para X_i . Cuando el modelo de medición es explícito, La cantidad del vector de salida es (ecuación 2.41) [15]:

$$y_r = f(x_r), r = 1, \dots, M \quad (2.41)$$

Las modificaciones necesarias son hechas si X_i no es independiente y más si una PDF conjunta es asignada para él. Estas modificaciones son implícitas, las cantidades de los vectores de salida so obtenidas de las soluciones de las ecuaciones (ecuación 2.42) [15]:

$$h(y_r, x_r) = 0, r = 1, \dots, M \quad (2.42)$$

Una representación discreta de la función de distribución para el vector de la cantidad de salida están formados por los valores M del vector de salida hallado en la ecuación 2.43. En general, Esta representación es una matriz G de dimensiones $m \times M$ donde la columna r -ésima es el r valor de vector de cantidad de salida. Para un modelo univariado, G es un vector de una sola dimensión [15].

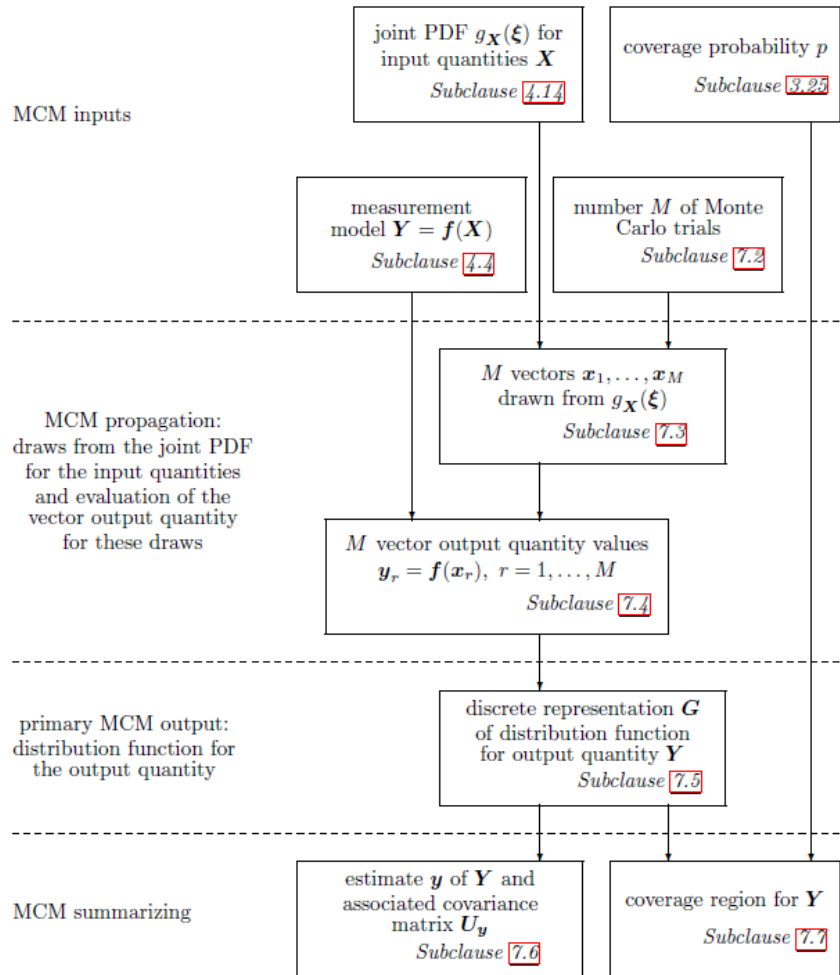
El promedio y la matriz de covarianza (ecuaciones 2.43 y 2.44):

$$\tilde{y} = \frac{1}{M} (y_1 + \cdots + y_M) \quad (2.43)$$

$$U_{\tilde{y}} = \frac{1}{M-1} [(y_1 - \tilde{y})(y_1 - \tilde{y})^T + \cdots + (y_M - \tilde{y})(y_M - \tilde{y})^T] \quad (2.44)$$

Son tomados, respectivamente, como un estimado y de Y y la matriz de covarianza U_y asociada a y [7].

Figura 2-5: Proceso MCM para cuantificar PDF para medición [7].



2.6 Teorema de la desigualdad de Chebyshev

Partiendo del hecho que un dato tomado en un momento k debe tener una probabilidad de estar dentro de los diferentes intervalos de los datos tomados anteriormente, es decir, entre intervalos definidos por las desviaciones estándar de un conjunto $K - 1$ anteriores a

este. Para efectos de esta tesis, un dato en un momento k tendrá la probabilidad de intervalo para $\pm 1\sigma, \pm 2\sigma$ y $\pm 3\sigma$ de los datos tomados anteriores a ese momento.

Según lo anterior, para una medida m tomada en momento k observado por un medidor m_i y σ de los valores del conjunto $K - 1$ anteriores, será:

$$Pr(m_k|m_i, \pm 1\sigma) + Pr(m_k|m_i, \pm 2\sigma) + Pr(m_k|m_i, \pm 3\sigma) = 1 \quad (2.46)$$

3. Metodología

Este capítulo contiene los pasos y requisitos necesarios para implementar una metodología conceptual de aumento de la certeza de la detección de desviaciones L/G utilizando la fusión de información de múltiples medidores de flujo en una línea de transporte de hidrocarburos refinados partiendo de una estrategia para la identificación por medidores interlínea y contabilización de pérdidas o ganancias por segmentos de línea. Se presentarán los pasos de manera ampliada y los requisitos mínimos deben cumplirse. Por tanto, la efectividad de la metodología está basada en la disciplina con la que se establezcan y ejecuten los requerimientos y los diferentes procesos propuestos.

Para evaluar la efectividad de esta metodología, se presentarán herramientas que permiten verificar implementación de la metodología. Estas se basarán en propuestas estadísticas clásicas e invariantes frente a las opciones de la metodología para garantizar la homogeneidad en la evaluación entre las mismas.

Algunas restricciones se presentarán a medida en que la propuesta del método se vaya desarrollando. Estas restricciones tienen como finalidad garantizar un desempeño adecuado de la metodología a las diferentes condiciones en una implementación posterior de dicha metodología. Se buscará entonces proponer estas restricciones claramente para que en el proceso posterior de implementación de la metodología puedan ser tenidas en cuenta y mitigadas para llegar a resultados óptimos.

3.1 Arquitectura conceptual

La metodología propuesta se integraría en un modelo de operativo de un Ducto de Transporte de Hidrocarburos (DTHR) de 5 capas:

- Adquisición, compensación y clasificación de la información de operación del DTHR.
- Comunicación y centralización de la información de DTHR.

- Fusión y Análisis de información de DTHR.
- Supervisión de la información fusionada y analizada de DTHR.
- Control y toma de decisiones basados en la información fusionada y analizada del DTHR.

La metodología que se propondrá en esta tesis se refiere y subscribe solo dentro de la tercera capa de este modelo operativo. Las condiciones y operatividad de las otras capas no hacen parte del desarrollo de esta metodología y se deben garantizar su óptima operación antes de considerar una implementación de la misma en un esquema operativo.

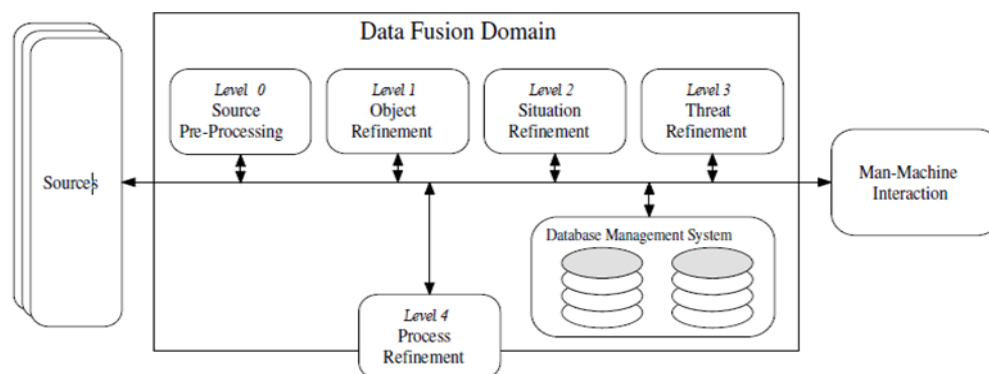
La metodología propuesta y suscrita dentro de la capa 3 del modelo operativo propuesto anteriormente. Dentro de esta capa se aplicara el modelo de fusión de información presentada por la “US Joint Directors of Laboratories (JDL)”. Esta fue propuesta en 1985 guiada por el Departamento de Defensa de dicho país.

El modelo JDL comprende cinco niveles de procesamiento y uno almacenamiento de datos que son interconectados por una información común. Esto quiere decir que los cinco niveles no tienen que ser procesados en estricto orden ya que pueden ser ejecutados una vez se necesite. Los elementos del modelo se describen a continuación [16]:

- **Fuentes:** Las fuentes proveen información de una variedad de fuentes de datos, como sensores, información a priori, bases de datos, entradas hechas por humanos.
- **Pre-procesamiento de fuentes (nivel 0):** Se busca con este elemento reducir la carga de procesamiento de la fusión de información por la valoración y calidad de la información a ser apropiadamente procesada.
- **Refinamiento de Objeto (nivel 1):** se realizan procesos de alineamiento (transformación de datos a referencias consistentes en valor y unidades), asociación (usando métodos de correlación), seguimiento e identificación usando métodos de clasificación.
- **Refinamiento de Situación (nivel 2):** Encuentra una descripción contextualizada de una relación entre los objetos y los eventos observados.

- **Refinamiento de Amenaza (nivel 3):** Basado en conocimiento previo y predicciones sobre una situación futura, este nivel describe las inferencias de las vulnerabilidades y oportunidades para la operación.
- **Refinamiento del Proceso (nivel 4):** Es un metaproceso que se encarga de monitorear la operación del sistema y relocaliza fuentes para alcanzar metas particulares.
- **Sistema de almacenamiento de datos:** La tarea de este sistema es monitorear, evaluar, adicionar, actualizar y proveer información para el proceso de fusión.
- **Interacción Hombre-Máquina:** Esta parte provee una interface para las entradas humanas y comunicaciones de los resultados de la fusión a los operadores y usuarios.

Figura 3-1: Modelo JDL para fusión de datos [16].

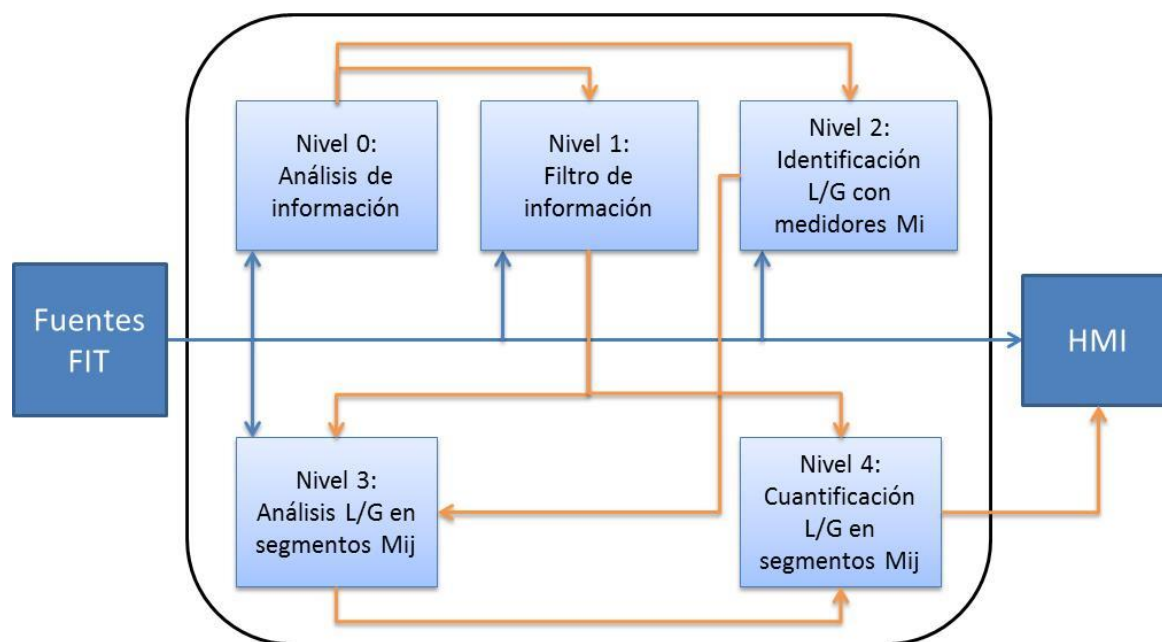


Para garantizar que las fuentes de la información sean adecuadas para la implementación de esta metodología a una operación real, los datos deben ser recolectados, compensados, sincronizados y etiquetados previamente por los sistemas de medición y adquisición (SCADA) de datos remotos con los cuales debe contar el sistema control del DTHR. Estos datos deben concentrarse en un punto para conservar la integridad y la no duplicidad de la misma. En conclusión, las fuentes serán evaluadas, verificadas y aseguradas previamente para una implementación o prueba de la misma metodología.

Para el caso particular de la metodología que se presentará en esta tesis, se aplicaran todos los niveles del modelo JDL para la fusión de información de un DTHR. Dado que

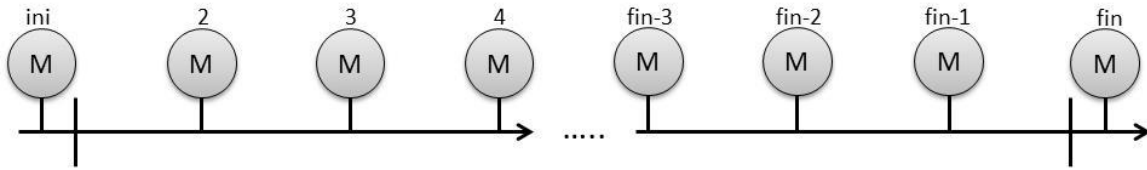
esta tesis estará enmarcada a nivel metodología la interacción hombre máquina y almacenamiento no se tendrá en cuenta ni la integridad de las fuentes. Los resultados de las simulaciones que se harán posteriormente mostrarían una potencial efectividad de la metodología si se llegará a implementar en una operación normal o contingencia de un DTHR.

Figura 3-2: Modelo para la integración de datos de un DTHR.



3.2 Arquitectura normalizada de un DTHR instrumentado

Para simplificar un modelo base se considerará una sola línea en la cual cuenta con un número de mediciones de flujo a lo largo de esta (m_i). Cada una de estas mediciones tiene asociada una incertidumbre técnica en la medición (σ_i) para cada medidor sobre la línea. El flujo medido en cada punto se encuentra en función de la medición directa, puede ser compensada en el mismo punto por presión y temperatura, y la incertidumbre técnica asociada a cada medidor $M_i = f(m_i, \sigma_i)$. Para una línea de transporte con diferentes puntos de medición se ilustraría de la siguiente figura 3.3.

Figura 3-3: Sistema simplificado de transporte de hidrocarburos por ductos.

Llevando este modelo al mundo real, la medición realizada en el inicio y fin del sistema tienen una incertidumbre menor a los medidores en la línea dada las consideraciones de espacio y seguridad necesaria para garantizar una medición adecuada para niveles de custodia de la operación. Se utilizan equipos de medición basados en turbinas con elementos fijos de volumen constante para probar sus constantes de medición. Estos equipos tendrían una desviación menor de 0,5%, conocidos como elementos de entrega en custodia. Para efectos prácticos de simulación y pruebas de funcionamiento que se harán en esta tesis, se utilizará un valor de 0,5% para la desviación de estos equipos.

Las mediciones en los puntos terminales son realizadas por equipos de entrega en custodia repetidamente calibrados contra probadores fijos de volumen constante, los equipos de medición interlínea no cuentan con estos probadores u otro equipo patrón para contraste. Por tanto, se buscará por medio de filtros reducir la incertidumbre de las mediciones hechas por los equipos interlíneas basados en las características de medición de custodia de los equipos terminales. Para efectos prácticos de simulación y pruebas de funcionamiento que se harán en esta tesis, se utilizará un valor de 5% para la desviación de equipos de medición de flujo entre línea.

3.2.1 Modelo de operación

Si el sistema DTHR se encuentra en estado estable, un llenado de línea constante y no hay pérdidas, es decir en un momento k :

$$M_{ini(k)} \cong M_{fin(k)} \quad (3.1)$$

Debería entonces:

$$M_{2(k)} \cong M_{3(k)} \cong \dots \cong M_{fin-2(k)} \cong M_{fin-1(k)} \quad (3.2)$$

Si suponemos que los equipos terminales tienen una menor incertidumbre (0,5%) que los no terminales o interlínea (5%), entonces no se podría decir para un momento k :

$$|M_{ini(k)} - M_{fin(k)}| \not\ll |M_{2(k)} - M_{3(k)}| \quad (3.3)$$

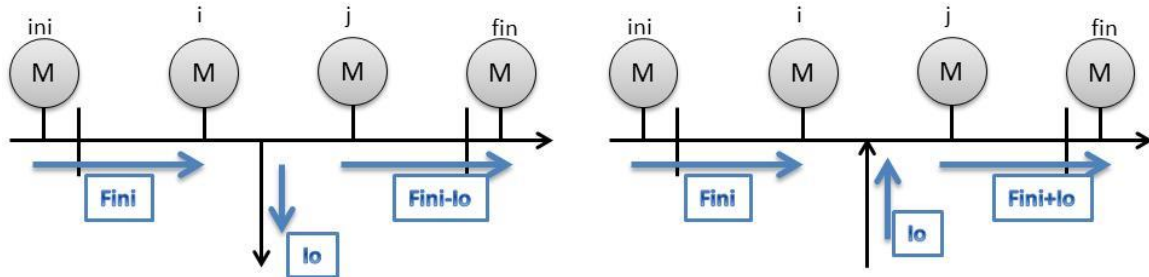
Esta diferencia resulta que para cada momento k , la incertidumbre de cada medidor, sea terminal o interlínea depende del factor de aleatoriedad de ese momento preciso. Por ejemplo, a pesar que los medidores terminales pueden encontrarse en un momento k en sus extremos de incertidumbre, esta diferencia puede ser mayor que la diferencia entre medidores línea que tengan incertidumbre cercana a cero en ese mismo momento k . Dada esta ambigüedad se debe enfatizar que el modelo de fusión sensor a aplicar debe llevarse bajo principios estadísticos y probabilísticos.

3.2.2 Modelo de desviación operativa por pérdida o ganancia (L/G)

Para los niveles de identificación y cuantificación de desviaciones L/G se propondría un modelo simplificado de un segmento de la línea. En este modelo se apoyará las interpretaciones numéricas que permitirán aplicar las diferentes teorías que se aplicarían.

Un modelo de L/G parte de la propuesta que existe un segmento cualquiera de la línea m_{ij} definido por los medidores m_i y m_j . En este segmento existe una desviación producida por un evento inherente o no a la operación de la línea, una pérdida de producto o una ganancia por una inyección de cualquier fluido.

Las L/G pueden presentarse por efectos térmicos menores pero la diferencial de velocidad de flujo en el ducto es baja y lenta. Para los casos de hurto, fugas parcial y/o total, las diferenciales de velocidad son más rápidas y con un mayor volumen de salida, lo que evidenciaría más claramente este tipo de desviación, se utilizará entonces esta consideración para analizar los efectos de la misma. En la gráfica 3.4 se mostrará el modelo de pérdida para el segmento m_{ij} .

Figura 3-4: Modelo de pérdida/ganancia para el análisis del segmento m_{ij} .

En un mismo momento k se puede tener modelos de ganancia y pérdida asociados a condiciones operativas de la línea. Por efectos de aumento de temperatura en un sector habría pérdidas del segmento m_{ij} , es semejante cuando la temperatura baja habrían ganancia en el sector. Estas variaciones dependen de la sección del ducto, longitud del segmento y de los gradientes de temperatura. Estos valores pueden ser despreciables si los gradientes son significativos o los valores de flujo son muy altos ya sobre un conjunto de tiempos K . Para el caso particular de esta tesis, esta aseveración se tomará en todos los casos de prueba teórica de la metodología.

3.3 Niveles metodológicos a implementar

A continuación presentaremos las estrategias que se utilizarán para cada nivel propuesto basado en el modelo JDL. Cada nivel tendrá un diferente tratamiento matemático y metodológico debido a notoria diferencia entre cada uno de estos niveles y sus resultados. Para un mejor entendimiento de cada uno de los procesos, se cambiarán los nombres de cada uno de estos pero se conservará el número nivel para mantener una clasificación fijada al modelo base.

Con el fin de reducir la pérdida en la calidad de los datos, es necesario que los datos no sean llevados a cálculos por unidad o ponderados. Este reproceso genera incertidumbre adicional por un reproceso adicional del mismo.

3.3.1 Filtro de la información (nivel 1)

Dado que la información observada por los medidores tiene un comportamiento aleatorio que dificulta su interpretación, es necesario tratar esta antes de fases posteriores de la metodología y llevarlas a niveles menores de desviación sin impactar negativamente sus características estadísticas. Se propondrá entonces la inclusión y aprovechamiento de herramientas estadísticas para filtrar la información siempre buscando tres premisas:

- Reducción de la incertidumbre de la muestra m_n para el momento k con respecto al valor más preciso del sistema en conjunto.
- Cambios menores en los valores de promedio del conjunto de los valores m_n para todos los momentos k .
- Cambios menores en la linealidad y aumento en la curtosis de la función de densidad de probabilidad.

Partiendo de estas tres premisas, se buscará implementar un filtro que tengan la menor incidencia en el valor totalizado de los valores para todos los momentos k . Se propondrá entonces dos filtros que facilitan dicha tarea, filtro Kalman y filtro Peso ponderado. Estos filtros pueden ser implementados desde la misma fase de adquisición de datos permitiendo descentralizar el procesamiento de estas señales en un caso tal que se requiera.

Dado que estos filtros incrementan la “masa” del sistema, se debe ser conservador en el momento de implementación ya que se puede afectar la respuesta a cambios operacionales del sistema. Al incrementar el factor q o A de cada filtro, este debe tomar iteraciones adicionales hasta converger en el comportamiento de las medidas. Estas demoras ocasionarían que la metodología pierda oportunidad y claridad en las respuestas posteriores por parte del operador del DTHR.

Filtro Kalman para datos de Entrada

Tomando como base teórica el numeral 2.4.1 de este documento, se podrá implementar el filtro Kalman a cada una de las mediciones que se tengan del DTHR. Este filtro nos permitirá analizar el comportamiento del flujo eliminando los valores más desviados del promedio del flujo pasante.

Al implementar el filtro Kalman se debe tener en cuenta que los datos que se van a filtrar tengan la integridad suficiente que reduzca posibles errores en una interpretación de los mismos. La información que entra al filtro debe ser procesada por el nivel 0.

Se parte de la suposición que el flujo para el tiempo de análisis es constante o con cambios muy pequeños en comparación con el flujo pasante, se asume que el modelo medido es constante, por lo tanto, $F = 1$ para todo $t \geq 0$.

Tomando las ecuaciones 2.15, 2.16 y 2.17, las variables de estado pueden ser reducidas a escalares dado que vamos a filtrar una única medida. Dado que los resultados de este filtro se darán en la misma escala que se entregan los valores de entrada, se tomará un $H = 1$.

La variable de control V_k no será utilizada dado que solo se adquirirá información y no se tendrá nivel de control de la misma. Para este modelo se asumirá que la variable R_{k+1} (ecuación 2.21) es igual que la desviación estándar de la medidas observadas por el medidor terminal. Dado que el proceso es escalar $P = p$ y el proceso no está bien definido se podrá tomar un $Q = q$, esto dependerá de la respuesta que se quiere del filtro (tabla 3.1).

Tabla 3-1: Proceso de filtro Kalman.

Entrada	Procesos	Salida
Valores de m_n para todo momento k . Valores de q y p , se propone un $q = 0.01$ y un $p = 1$. $r = s^2(m_n)$ Para el medidor m_n .	Para cada momento k : $x_{k k-1} = m_{n_{k-1}}$ $p_{k k-1} = p_{k-1 k-1} + q$ $K_k = p_{k k-1}(p_{k k-1} + r)^{-1}$ $p_{k k} = (1 - K_k) p_{k k-1}$ El valor de filtrado m_n es $\boxed{m_n} = x_{k k} = x_{k k-1} + K_k(m_{n_k} - x_{k k-1})$	Valores de m_n para todo momento k filtrados $\boxed{m_n}$.

Filtro de peso ponderado para datos de Entrada

Es filtro se basa en la presunción de veracidad de un valor tomado contrastándolo con un valor actual por un peso A que varía entre 0 y 1. Este peso es tomado dependiendo de la respuesta que se quiere del filtro y es de amplio uso en sistemas control de ruidosos por su fácil implementación aunque por su simplicidad no permite modificaciones que amplíen el rango de utilización.

Hay que tener en cuenta que la independencia de la medición es necesaria para este filtro, tanto como para el filtro Kalman. Este filtro permite balancear las cargas del dato anterior con el actual y diferencias no lineales dentro de la misma medición lo pueden “desorientar” o perder la medida. Se debe asegurar igualmente la integridad de los datos que serán procesados.

Tabla 3-2: Proceso de filtro de peso ponderado.

Entrada	Procesos	Salida
Valores de m_n para todo momento k . Valor de A .	Al iniciar el proceso de filtrado $B = 1 - A$ Para cada momento k : $\boxed{m_n} = A \cdot m_{n_k} + B \cdot m_{n_{k-1}}$	Valores de m_n para todo momento k filtrados $\boxed{m_n}$.

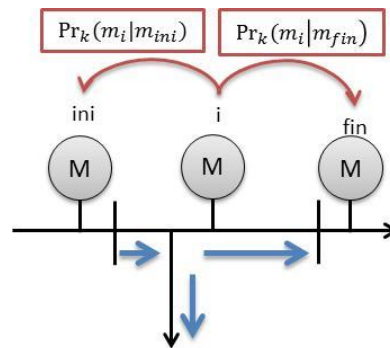
3.3.2 Identificación de desviaciones operativas L/G (nivel 2)

Se procederá con antelación con el cálculo de las probabilidades de los valores de los medidores en el momento k en función de los valores tomados en los bloques de valores $K - 1$ de los otros medidores según el teorema de Chebyshev. Los indicadores de probabilidad tendrán la notación $\Pr_k(m_m|m_n)$ donde los subíndices m y n asignan el número de medidor y estima la probabilidad de que el valor del medidor m_m en un momento k se encuentre entre seis desviaciones estándar ($\pm 3\sigma$) definido por todos los valores de m_n en el conjunto $K - 1$. Se puede reducir el intervalo de desviaciones para

una mejor respuesta del sistema ya que la probabilidad intersección de los datos sería menor para el momento k en el conjunto $K - 1$ pero induciría a mayores errores de estimación. En la figura 3.5 se presentará en el modelo de pérdida/ganancia como se originarían las probabilidades para este proceso.

El uso del conjunto de valores $K - 1$ permite que el sistema se adapte a las condiciones reales de operación en todo momento. Esta propiedad permite a esta metodología integrarse a condiciones de cambios normales operativos inherentes a la operación sin que se pierda la objetividad en la medición de desviaciones de pérdida o ganancia anormales dentro del proceso. Estos cambios se pueden enmarcar en los porcentajes de sensibilidad al sistema por efectos de desviaciones de L/G.

Figura 3-5: Análisis de probabilidad para los medidores del segmento m_{ij} .



Dado que la información debe revisarse por bloques de datos, se debe asignar un valor de esta estimación a iniciar. Estos tamaños dependen directamente de la variabilidad del proceso de transporte tales como las propiedades del producto y temperatura ambiente. La estrategia de análisis de este documento se dará por bloques de datos $K - 1$, donde K está definido como un vector de bloque de datos para todos los momentos k previos en ese bloque.

La estrategia planteada será la obtención del valor de la probabilidad de intersección de las mediciones contra la medición completamente anterior para el caso de los medidores internos del DTHR, los medidores inicial y final no se analizarían pero servirán de soporte para la verificación de los medidores internos. Un análisis de intersección de datos contra valores de medidor anterior e inicial o final (m_{ini} y m_{fin}) dada la posición de que el medidor se encuentre en la línea.

3.3.3 Aislamiento de desviaciones L/G en segmentos m_{ij} (nivel 3)

Se busca con este análisis generar indicadores de comportamiento de las L/G de la línea basados en los medidores de flujo que se encuentran instaladas en ella. Tomando las probabilidades estimadas en la fase de identificación de L/G por segmentos, estas se analizarán y fusionarán con valores de otros detectores adyacentes para generar información adicional para descartar fallas o situaciones inherentes a la operación de cada medidor.

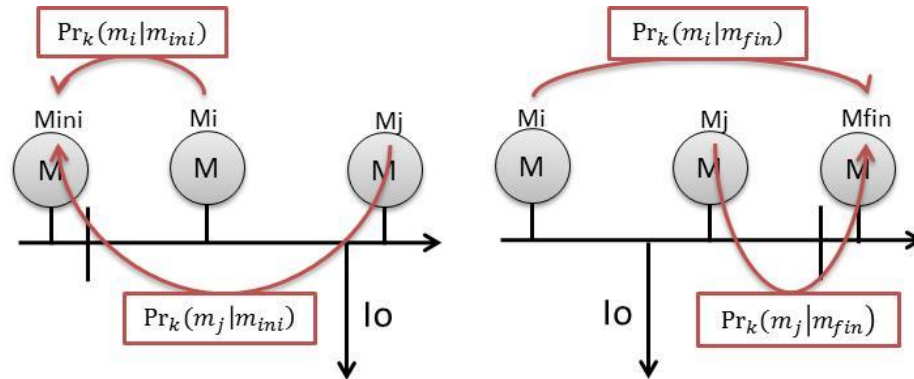
Tabla 3-3: Proceso de análisis de probabilidad de intersección para medidores.

Entrada	Procesos	Salida
Valores de m_i, m_{ini} y m_{fin} para los bloques de datos $K - 1$. Valores de probabilidad de operación norma previas $Pr_k(m_{io} m_{inio})$ $Pr_k(m_{io} m_{fino})$ Para cada medidor m_i .	Para cada medidor m_{ij} para bloques $K - 1$: $Pr_k(m_i m_{ini}) = Pr(\overline{m_i} m_{ini})$ $Pr_k(m_i m_{fin}) = Pr(\overline{m_i} m_{fin})$ Para todas las probabilidades $Pr_k(m_j m_{ini})$, $Pr_k(m_j m_{fin})$: $\begin{cases} Pr_k(m_i m_{ini}) \geq +Pr_k(m_{io} m_{inio}) \therefore P_{i,ini}[k, 1] = 1 \\ Pr_k(m_i m_{ini}) \leq -Pr_k(m_{io} m_{inio}) \therefore P_{i,ini}[k, 1] = 0 \\ Pr_k(m_i m_{fin}) \geq +Pr_k(m_{io} m_{fino}) \therefore P_{i,fin}[k, 1] = 1 \\ Pr_k(m_i m_{fin}) \leq -Pr_k(m_{io} m_{fino}) \therefore P_{i,fin}[k, 1] = 0 \end{cases}$	Para momento k : $Pr_k(m_j m_{ini})$, $Pr_k(m_j m_{fin})$ Periodo de L/G: $T_{(L / G)}$

En un caso particular, este análisis y/o fusión de la información observada dará indicadores de probabilidad para cada medidor que al ser contrastados con los valores normales de ducto, analizados previamente, indicarán una desviación de ganancia o pérdida en el DTHR por cada segmento m_{ij} . Este caso permitirá cuantificar las L/G en un momento k sumado a la fusión de segmentos que se hará en el nivel 4.

Se harán dos análisis para las situaciones en que los datos se interceptan con los valores del medidor inicial m_{ini} y final m_{fin} . Los valores analizados a partir del medidor inicial permitirán analizar desde que segmento se da la pérdida y los valores tomados desde el medidor final definirán los periodos de pérdida dentro de los valores.

Figura 3-6: Análisis de probabilidad para observaciones de los medidores T.



Fusión de datos del segmento m_{ij} por Dempster-Shafer

Se puede partir de teoría de Dempster-Shafer donde la función de credibilidad (ecuación 2.27) da las bases de la fusión de información. Esta función está dada por la sumatoria de los pesos por el segmento del subgrupo a la que ella pertenece. Se puede entonces proceder a generar un indicador para el momento k de los medidores m_i y m_j fusionándolos. Si llevamos este procedimiento a todos los medidores de línea, este indicador permitirá analizar para el momento k el estado de todos los medidores que se encuentran instalados a lo largo de la línea y contrastarlo con los niveles de operación previos a este momento (tabla 3.4).

Fusión de datos del segmento m_{ij} por Bayes

Se podrá optar también por el uso de la teoría de Bayes para fusionar los datos por segmentos. Se utilizará una estrategia similar a la usada por el método Dempster-Shafer iniciando un análisis de la intersección de los valores k momento en función del bloque $K - 1$. Este método reduciría la carga computacional para el análisis pero puede llegar a ser menos sensible dada su simplicidad de aplicación. Utilizaremos para el análisis solo la información tomada de la primera desviación dado que esta cuenta con certeza que es la más cercana al valor real en que se encuentra el flujo por el DTHR (tabla 3.5).

Tabla 3-4: Fusión de datos por segmentos m_{ij} por Dempster-Shafer.

Entrada	Procesos	Salida
Valores de m_i, m_j, m_{ini} y m_{fin} para los bloques de datos $K - 1$.	<p>Para cada segmento m_{ij} para bloques $K - 1$ y utilizando el Teorema de la desigualdad de Chebyshev:</p> $\Pr_k(m_i m_{ini}, \pm 1\sigma), \Pr_k(m_i m_{fin}, \pm 1\sigma)$ $\Pr_k(m_i m_{ini}, \pm 2\sigma), \Pr_k(m_i m_{fin}, \pm 2\sigma)$ $\Pr_k(m_i m_{ini}, \pm 3\sigma), \Pr_k(m_i m_{fin}, \pm 3\sigma)$ $\Pr_k(m_j m_{ini}, \pm 1\sigma), \Pr_k(m_j m_{fin}, \pm 1\sigma)$ $\Pr_k(m_j m_{ini}, \pm 2\sigma), \Pr_k(m_j m_{fin}, \pm 2\sigma)$ $\Pr_k(m_j m_{ini}, \pm 3\sigma), \Pr_k(m_j m_{fin}, \pm 3\sigma)$ <p>Fusionando bajo a m_i y m_j bajo la regla de la combinación de Dempster-Shafer (ecuación 31). Para análisis desde las probabilidades de intersección con valores del medidor inicial m_{ini}:</p> $\widehat{m}_{ij_{ini}} = \Pr_k(m_i m_{ini}, \pm 1\sigma) * \Pr_k(m_j m_{ini}, \pm 1\sigma)$ $+ \Pr_k(m_i m_{ini}, \pm 1\sigma) * \Pr_k(m_j m_{ini}, \pm 2\sigma)$ $+ \Pr_k(m_i m_{ini}, \pm 1\sigma) * \Pr_k(m_j m_{ini}, \pm 3\sigma)$ $+ \Pr_k(m_i m_{ini}, \pm 2\sigma) * \Pr_k(m_j m_{ini}, \pm 1\sigma)$ $+ \Pr_k(m_i m_{ini}, \pm 3\sigma) * \Pr_k(m_j m_{ini}, \pm 1\sigma)$ <p>Para análisis desde las probabilidades de intersección con valores del medidor inicial m_{fin}:</p> $\widehat{m}_{ij_{fin}} = \Pr_k(m_i m_{fin}, \pm 1\sigma) * \Pr_k(m_j m_{fin}, \pm 1\sigma)$ $+ \Pr_k(m_i m_{fin}, \pm 1\sigma) * \Pr_k(m_j m_{fin}, \pm 2\sigma)$ $+ \Pr_k(m_i m_{fin}, \pm 1\sigma) * \Pr_k(m_j m_{fin}, \pm 3\sigma)$ $+ \Pr_k(m_i m_{fin}, \pm 2\sigma) * \Pr_k(m_j m_{fin}, \pm 1\sigma)$ $+ \Pr_k(m_i m_{fin}, \pm 3\sigma) * \Pr_k(m_j m_{fin}, \pm 1\sigma)$	Vector de probabilidad segmento m_{ij}

Tabla 3-5: Fusión de datos por segmentos m_{ij} por Bayes.

Entrada	Procesos	Salida
Valores de m_i, m_j, m_{ini} y m_{fin} para los bloques de datos $K - 1$.	<p>Para cada segmento m_{ij} para bloques $K - 1$ y utilizando el Teorema de la desigualdad de Chebyshev:</p> $\Pr_k(m_i m_{ini}, \pm 1\sigma), \Pr_k(m_i m_{fin}, \pm 1\sigma)$ $\Pr_k(m_j m_{ini}, \pm 1\sigma), \Pr_k(m_j m_{fin}, \pm 1\sigma)$ <p>Fusionando bajo a m_i y m_j bajo la teoría de Bayes (ecuación 31). Para análisis desde las probabilidades de intersección con valores del medidor inicial m_{ini}:</p> $A_{ij} = \Pr_k(m_i m_{ini}, \pm 1\sigma) * \Pr_k(m_j m_{ini}, \pm 1\sigma)$ $B_{ij} = * \Pr_{k-1}(m_i m_{ini}, \pm 1\sigma)^{-1} \Pr_{k-1}(m_j m_{ini}, \pm 1\sigma)^{-1}$ $\widehat{m}_{ij_{ini}} = A_{ij} * (m_i * m_j)_{k-1} * B_{ij}^{-1}$ <p>Para análisis desde las probabilidades de intersección con valores del medidor final m_{fin}:</p> $A_{ij} = \Pr_k(m_i m_{fin}, \pm 1\sigma) * \Pr_k(m_j m_{fin}, \pm 1\sigma)$ $B_{ij} = * \Pr_{k-1}(m_i m_{fin}, \pm 1\sigma)^{-1} \Pr_{k-1}(m_j m_{fin}, \pm 1\sigma)^{-1}$ $\widehat{m}_{ij_{fin}} = A_{ij} * (m_i * m_j)_{k-1} * B_{ij}^{-1}$	Vector de probabilidad segmento m_{ij}

Estos mismos procesos pueden ser llevados para fusionar los datos de 2σ y 3σ . Se deben proceder de manera independiente para cada uno. La interpretación de estos datos daría como resultado como una medida única para cada segmento en vez de tener dos valores dados independientemente de dos medidores.

3.3.4 Caracterización de una desviación operativa L/G (nivel 4)

Partiendo de la información obtenida por los niveles 2 y 3 se podrá hacer una cuantificación estimada de la pérdida/ganancia para la información de los diferentes segmentos m_{ij} de un DTHR. Esta cuantificación dada que es una medida tendrá una incertidumbre

asociada a la estrategia de análisis a tomar para la fusión y combinación de la información analizada por los niveles anteriores.

Contrastando la información del nivel 2 de la identificación de una pérdida determinada por uno de los medidores de la línea comparada contra los medidores adyacentes e inicial y final contra la información del nivel 3 de los periodos de pérdida/ganancia (L/G) y segmento asociado analizaremos alguna opciones posibles a tomar para una pérdida en un segmento m_{ij} para el periodo definido (L/G):

- Resta aritmética de las mediciones observadas sin filtrar.
- Resta aritmética de las mediciones observadas con filtro Kalman.
- Fusionar los valores observados sin filtrar de los segmentos definidos por un a L/G , es decir m_{ini} hasta i y m_j hasta fin

Las alternativas 1 y 2 para el análisis difieren del uso de los filtros para mejorar la confiabilidad de los datos de entrada a la fusión de la información de los medidores. Estos se podrán hacer independientemente de la implementación de la fusión de los medidores utilizando los métodos expuestos para el nivel 1 (Kalman o ponderado).

Para la fusión de múltiples medidores podremos utilizar los métodos mencionados en el nivel 3. Se utilizará únicamente el caso de intersección $\alpha\sigma$ que está mencionado en el marco teórico de esta tesis debido a que la fusión por Bayes y Dempster-Shafer es solo usado para medición de probabilidades para el propósito de esta tesis.

Fusión de datos de los segmentos m_{ij} por promedio ponderado de confianza para mediciones m_n

Una alternativa óptima para la fusión de esta información que puede ser aplicada dada su baja carga computacional y bajo nivel de interpretación de valores sería utilizar el método del promedio ponderado de confianza. Se utilizarían los mismos datos de entrada al sistema y salidas de los medidores filtrados o sin filtrar.

Este método es bastante utilizado en diferentes aplicaciones de fusión de sensores dada la su característica de baja complejidad. Para efectos de esta tesis, se probará solo para fusionar información flujo de medidores directamente y no de probabilidades. Esto

permitirá mejorar la incertidumbre de las mediciones asociadas a las L/G en un periodo $T_{(L/G)}$.

Tabla 3-6: Fusión de datos por segmentos m_{ij} por promedio ponderados.

Entrada	Procesos	Salida
Valores de m_i, σ_i, m_j y σ_j para el momento k .	<p>Para cada segmento m_{ij}:</p> $w_i = (\sigma_i^2 * \sigma_j^{-2})^{-1}$ $w_j = (\sigma_j^2 * \sigma_i^{-2})^{-1}$ <p>Se procede a fusionar según la ecuación 36 y 37:</p> $\widehat{m}_{ij_k} = \sum_{a=1}^N m_a w_a = m_{i_k} w_i + m_{j_k} w_j$ $\sigma_{ij}^2 = \sum_{a=1}^N w_a^2 \sigma_a^2 = w_i^2 \sigma_i^2 + w_j^2 \sigma_j^2$	Vector de medición de L/G por segmentos $m_{ini,i}$ a $m_{j,fin}$

4.Evaluación Teórica

Para demostrar la operatividad y rendimiento de la metodología propuesta en el anterior capítulo, se llevará ésta a prueba en modelos DTHR simplificados. Esta simplificación partirá de un flujo constante de 100 unidades por periodo (u/p) en el medidor inicial y variará a medida que en diferentes sectores se induzcan pérdidas o ganancias de varias magnitudes.

Dada la poca variabilidad y “damping” de los controles de flujo que gobiernan este en la estación de bombeo, se tendrá un valor esperado (μ) normalizado de 100. Este valor permitirá visualizar el comportamiento de las diferentes situaciones de pérdida o ganancia de la línea.

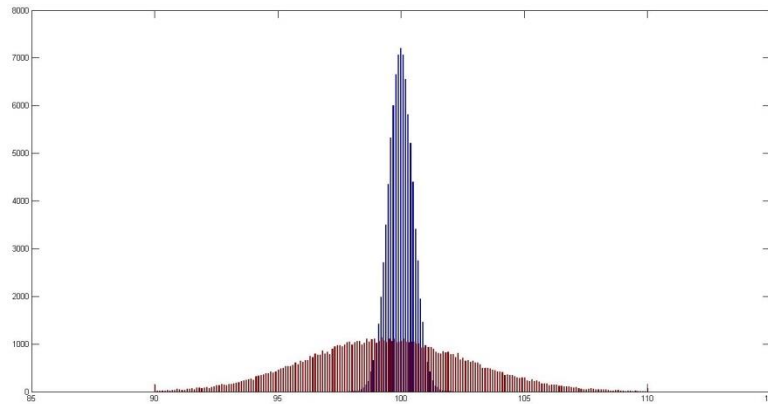
Para fines prácticos de evaluación y basados en el teorema central del límite, donde en términos generales el resultado de la suma de variables aleatorias tiende agruparse en distribuciones de probabilidad normales, se tomarán esta distribución para los medidores terminales T para un valor esperado (μ) de 100 y una desviación estándar (σ) de 0,005. Para los medidores interlinea I la distribución de probabilidad es modelada de la suma ponderada de distribuciones normales, uniformes y geométricas para un valor esperado (μ) de 100 y una desviación estándar (σ) de 0,05. Este modelo fue tomado de los datos de medidores de flujo ultrasónicos de doble haz con corrección volumétrica por temperatura instalado actualmente sobre líneas de trasiego. La información de soporte es confidencial dado que los datos son entregados por una empresa privada. En la figura 4.1 se puede observar la distribución de los datos entre los medidores T e I .

Este modelo de evaluación no se llevará por todas las fases de la metodología propuesta en el capítulo anterior y referente al modelo JDL. Se obviará el nivel 0 de análisis de información ya que los datos que se están generando para la evaluación están completos y son lógicos para estas pruebas. Para modelos reales que serán revisados

posteriormente y fuera del alcance de esta tesis, el nivel 0 aportará en robustecimiento del análisis de L/G de un DTHR.

El nivel 1 de filtrado de información se llevará no solo como un prerequisite para los otros niveles, si no que podrá ser utilizado en análisis posteriores de los resultados de estos niveles. Los niveles 2 y 3 se evaluarán en función sensibilidad del aislamiento y detección de L/G inducida en la prueba. Finalmente, el nivel 4 se analizará la desviación de medición de la L/G dependiendo de los diferentes arreglos que se proponen para el DTHR.

Figura 4-1: Diferencia de PDF entre medidores terminales e interlínea.



4.1 Probabilidad de mediciones de medidores con diferente incertidumbre.

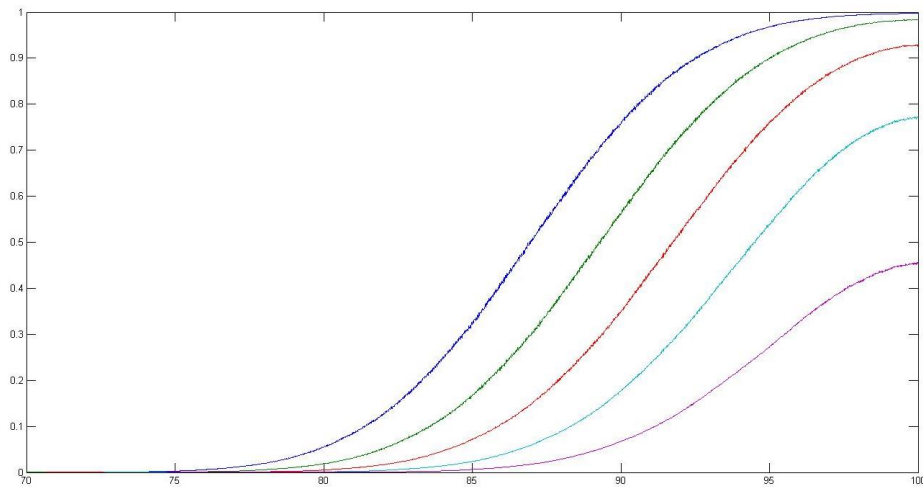
Para realizar esta identificación partimos de la suposición que el flujo entre los medidores m_i y m_j del DTHR es constante en el tiempo para un momento k determinado y este define el segmento m_{ij} de la línea. Esta identificación de pérdidas/ganancias planteada a continuación es solo posible, si y solo si para bloques de datos completos y es necesario aplicar la metodología del nivel 0 para esto.

Las mediciones en el segmento m_{ij} deben tener una relación estadística dada por la probabilidad que ambas mediciones se encuentren dentro un hiperelipse envolvente definida por la n-varianza del medidor menor contrastado del segmento en el bloque de datos completo. Esta varianza debe ser adquirida previamente y se pueden utilizar los

métodos planteados en el marco teórico en la sección 5 con el método de Monte Carlo y la evaluación tipo A de la incertidumbre.

Para un momento k , el valor m_j debe estar dentro de la función de distribución de probabilidad de m_i si la varianza de $\sigma_{m_i} \geq \sigma_{m_j}$ para un valor de F definido para el segmento. Por ejemplo, Para las condiciones anteriores y definiendo una desviación estándar del 5% de F para los medidores m_i y m_j con distribución normal, su probabilidad de intersección es de aproximadamente 99,73%, si existe una desviación del 10% del valor promedio de m_j se tendría una probabilidad de intersección de 75,63%. La relación entre estas intercepciones está presentada en la figura 4.2.

Figura 4-2: Probabilidad de mediciones entre un medidor interlínea y terminal

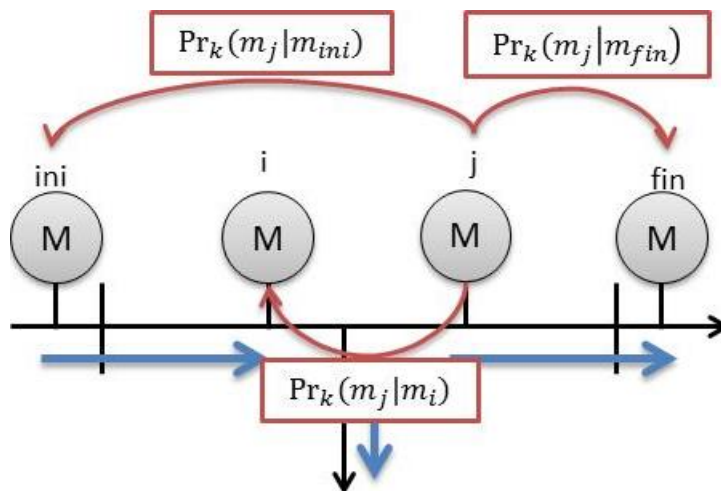


Entre más precisos sean los equipos de medición las L/G se pueden detectar a ratas de flujo utilizando este método. Por lo tanto, el uso de filtro de información como el Kalman o peso ponderado se recomienda para este análisis. El uso de estos filtros agruparan la información agrupada en los muestras del conjunto K . El tamaño de un conjunto K se definirá dependiendo de la cantidad de información de los medidores en periodos de tiempo en donde las variables físicas de la operación de un DTHR sean homogéneas. Estos grupos pueden ser definidos empíricamente, en DTHR analizados la variación es mayor a 10 minutos, para el caso concreto de este análisis, se tomarán periodos de 5 minutos con frecuencia de 4 muestras por segundo, por tanto un grupo $K = 1200$

muestras. Para reducir la carga computacional de los análisis se tendrá un grupo $K = 1000$ muestras.

Se procederá con antelación con el cálculo de las probabilidades de los valores de los medidores en el momento k en función de los valores tomados en los bloques de valores $K - 1$ de los otros medidores. Los indicadores de probabilidad tendrán la notación $Pr_k(m_m|m_n)$ donde los subíndices m y n asignan el número de medidor y estima la probabilidad de que el valor del medidor m_m en un momento k se encuentre entre seis desviaciones estándar ($\pm 3\sigma$) definido por todos los valores de m_n en el conjunto $K - 1$. Se puede reducir el intervalo de desviaciones para una mejor respuesta del sistema ya que la intersección de los datos sería menor en para el momento k en el conjunto $K - 1$ pero induciría a mayores errores de estimación. En la figura 4.3 se mostrará en el modelo de pérdida/ganancia como se originarían las probabilidades para este proceso.

Figura 4-3: Análisis de probabilidad para los medidores del segmento m_{ij} .



Según la propuesta anterior, para el medidor m_2 el análisis de $Pr_k(m_2|m_{ini})$ y sería igual para los dos indicadores. Esto aumentaría la capacidad del indicador al analizar este segmento dado el bajo nivel de incertidumbre del medidor inicial.

El uso del conjunto de valores $K - 1$ permite que el sistema se adapte a las condiciones reales de operación en todo momento. Esta propiedad permite a esta metodología integrarse a condiciones de cambios normales operativos inherentes a la operación sin que se pierda la objetividad en la medición de desviaciones de pérdida o ganancia

anormales dentro del proceso. Estos cambios se pueden enmarcar en los porcentajes de sensibilidad al sistema por efectos de desviaciones de L/G.

4.2 Evaluación de filtros de información (nivel 1)

Para evaluar el rendimiento de los filtros Kalman y Ponderado en función de las necesidades de la metodología se buscará la modificación de las variables de entrada para cada filtro. Se buscará una reducción de la dispersión de los valores de los medidores Interlinea dado que los medidores Terminales funcionarán como referencia confiable para toda la metodología.

Se modificará para el filtro Kalman el valor q y para el Ponderado el valor A . Los valores suministrados por la modificación de ambas variables suministrarán la interrelación que podrá tener este filtro para su uso en los siguientes niveles de la metodología. Es decir, si para un valor dado de q o A la respuesta es una serie de datos muy filtrado, la desviación estándar de los datos filtrados de un medidor interlinea es muy cercana a la de un medidor Terminal, este factor nos servirá para aumentar la precisión de la cuantificación en el nivel 4 pero no nos será de ayuda para la detección o aislamiento para el nivel 2 y 3 respectivamente ya que perdemos sensibilidad de ambos niveles.

Los valores filtrados serán evaluados y comparados en función de sus características estadísticas. Valores de promedio, límites máximos/mínimos, desviación estándar, varianza, curtosis serán evaluados para cada variación de las variables. Estas características nos brindan un marco objetivo de cómo se comporta los datos de medidores de flujo interlinea para un posterior aplicación en pasos de esta misma metodología.

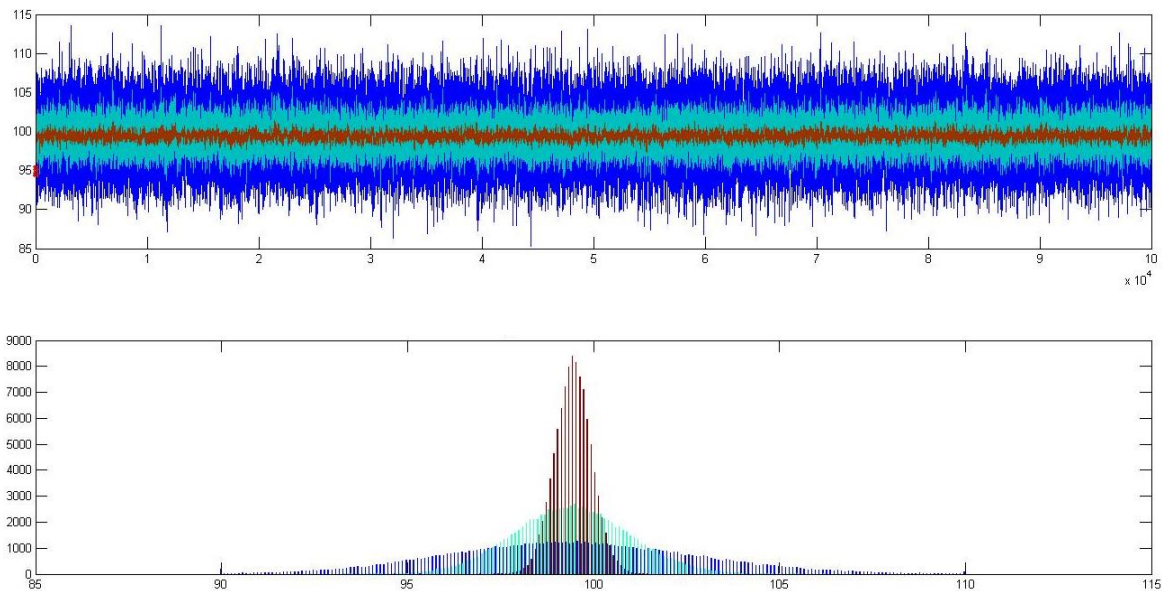
4.2.1 Evaluación de Filtro Kalman para mediciones Interlínea

Generando un valor aleatorio y modelado con información analizada de medidores reales de flujo interlinea se procedió a filtrar esta información por un filtro Kalman con diferentes valores de covarianza q diferentes. Se parte de un valor alto de 0.1 permitido por la literatura para este tipo de evaluaciones. En la tabla 1 se podrá identificar la respuesta en valores estadísticos de las señales una vez se use el filtro en estas. La columna con valor 0 es la caracterización de la señal de entrada sin modificación por el filtro.

Tabla 4-1: Filtro Kalman para varios factores de varianza q

q	0	0,1	0,05	0,01	0,005	0,001	0,0005	0,0001
Promedio	99,4061	99,4061	99,4061	99,4061	99,4061	99,4062	99,4062	99,4062
Máximo	113,4384	109,927	108,9487	106,1529	105,4185	103,3795	102,594	101,7963
Mínimo	85,4178	88,2948	89,6521	91,5894	92,8308	95,6084	96,1848	97,418
Desv.Est.	3,3008	2,5105	2,2108	1,5476	1,3108	0,8837	0,7449	0,5016
Varianza	10,8953	6,3028	4,8876	2,3951	1,7183	0,781	0,5549	0,2516
Curtosis	2,9568	2,9713	2,9812	3,0152	3,0338	3,0782	3,0881	3,0723

Analizando la información anterior, los valores de covarianza del filtro Kalman permiten filtrar con un buen rendimiento de las señales del medidor de flujo interlinea. Esto se observa en la poca o nula variación del promedio del valor crudo ($q = 0$) contra el valor de mínima covarianza ($q = 0,0001$). Reforzando la aseveración inicial de este párrafo, los rangos de los datos máximo y mínimo se ven reducidos en cada cambio en el valor de la covarianza repercutiendo igualmente en la desviación estándar y varianza de la misma. La curtosis permanece inalterada pese a que el apuntalamiento de la distribución es evidente.

Figura 4-4: Filtro Kalman para mediciones de flujo Interlínea¹.

Se puede notar que en los valores cuando la covarianza es ($q = 0,0001$) la desviación estándar es similar a la de un medidor de flujo Terminal. Este valor de covarianza será el

¹ Azules la señal cruda, Turquesa con covarianza de 0.05 y Rojo con una covarianza de 0.0001.

valor máximo utilizado en este análisis ya que un sentido lógico un valor de un medidor de flujo de alta incertidumbre no podría tener una menor incertidumbre de un medidor de flujo de baja incertidumbre por el uso del filtro ya que la realidad está observada por el medidor de baja incertidumbre y referenciada contra patrones fijos. En la figura 4.4 se puede visualizar el impacto del uso del filtro Kalman en los datos de los medidores de flujo interlinea.

Filtro Kalman para mediciones Interlinea con variaciones a bajas frecuencias

Para analizar el comportamiento de variaciones inducidas por el DTHR de baja frecuencia como el efecto de la temperatura ambiente en la línea se procede a incluir un valor sinusoidal de un pico máximo del 5% de los datos si no se tuviere esta variación. El valor de 5% es un valor exagerado del comportamiento real de una tubería en función de cambios del volumen almacenado en ella, se tomó este valor para probar en condiciones extrema la respuesta del mismo en frecuencias bajas.

Tabla 4-2: Filtro Kalman para varios factores de varianza q a bajas frecuencias.

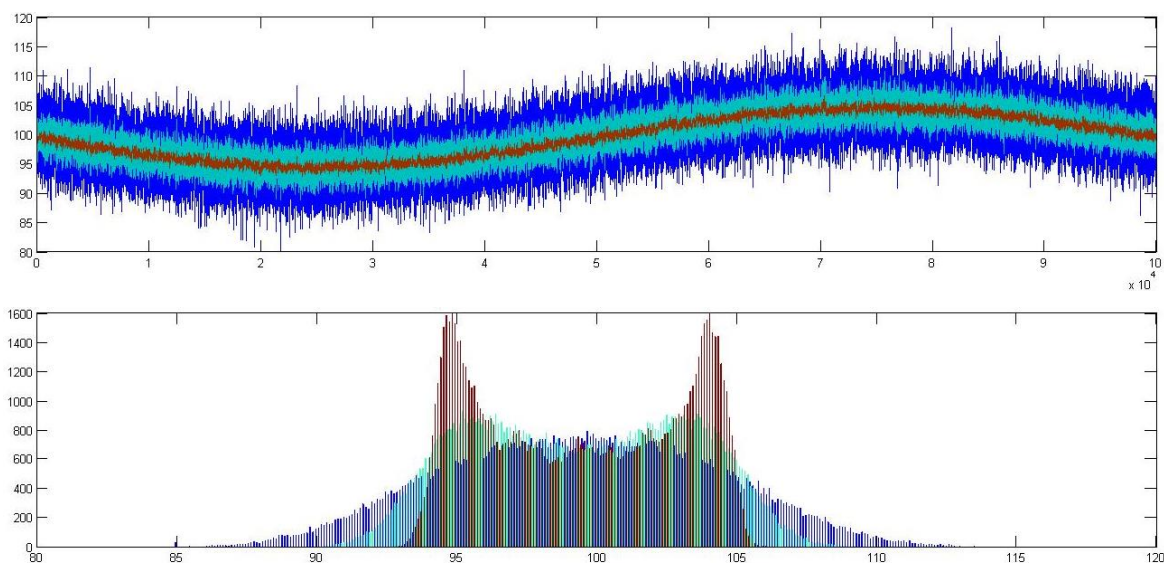
q	0	0,1	0,05	0,01	0,005	0,001	0,0005	0,0001
Promedio	99,4346	99,4345	99,4345	99,4345	99,4346	99,4346	99,4346	99,4345
Máximo	118,0531	114,0745	113,1353	110,6735	109,853	107,8582	107,2031	106,5422
Mínimo	80,0826	83,6045	85,1407	88,7658	89,7301	91,3446	91,8832	92,7743
Desv.Est	4,8609	4,3572	4,1903	3,8801	3,7919	3,6669	3,6361	3,5941
Varianza	23,6284	18,9852	17,559	15,0552	14,3788	13,4463	13,2213	12,9174
Curtosis	2,5562	2,3285	2,2174	1,9357	1,8326	1,6648	1,6195	1,5552

Como paso en la evaluación hecha para el flujo constante (4.2.1), la variación es casi nula para el promedio, lo que implicaría un comportamiento aceptable en términos de afectación a la medición. Las variaciones del 5% afectan sustancialmente los límites máximos y mínimos pero dado que es un error sistemático inducido se puede comparar con las otras medidas dando como resultado una variación importante, más de 12% para las covarianzas 0 y 0,0001. La afectación en la varianza y desviación estándar está ligada a los límites pero en función del análisis de medición para esta variación no se presta para un análisis pleno. La curtosis es claramente afectada por el hecho que el apuntalamiento para un incertidumbre del 5% es mucho mayor que para las desviaciones de 0,5% de la misma medida filtrada con una covarianza de 0,0001. Ver la figura 4.5 para una mejor observación del fenómeno.

Se puede notar en la gráfica asociada a los cambios de baja frecuencia que existe una gran cantidad de datos acumulados en las variaciones en los rangos +5% y -5%. Esto es importante en la medida que estas variaciones no son notorias para un medidor interlínea que no haya sido filtrado. Es una importante conclusión que se saca de este análisis ya que como se dijo anteriormente esta es una suposición extrema y las variaciones en la operación real son mucho más pequeñas rondando el 1%.

Se debe evaluar la operatividad del filtro a cambios repentinos en del flujo para condiciones operativas normales o adversas a la este. Los cambios rápidos en pérdidas de un DTHR se pueden clasificar como rompimientos parciales o totales, drenajes o hurtos. Por parte de las ganancias, se pueden clasificar como reinyecciones por parte de las planta adheridas al sistema aunque son condiciones no comunes en una operación.

Figura 4-5: Filtro Kalman para mediciones de flujo Interlinea a bajas frecuencias².



Filtro Kalman para mediciones Interlinea con variaciones rápidas del flujo

Dado el marco presentado, el filtro Kalman debe asegurar una respuesta rápida ante estos cambios para que las señales observadas y filtradas describan lo más preciso la realidad del flujo pasante. Una respuesta lenta del filtro ocasionaría una subvaloración

² Azules la señal cruda, Turquesa con covarianza de 0.05 y Rojo con una covarianza de 0.0001.

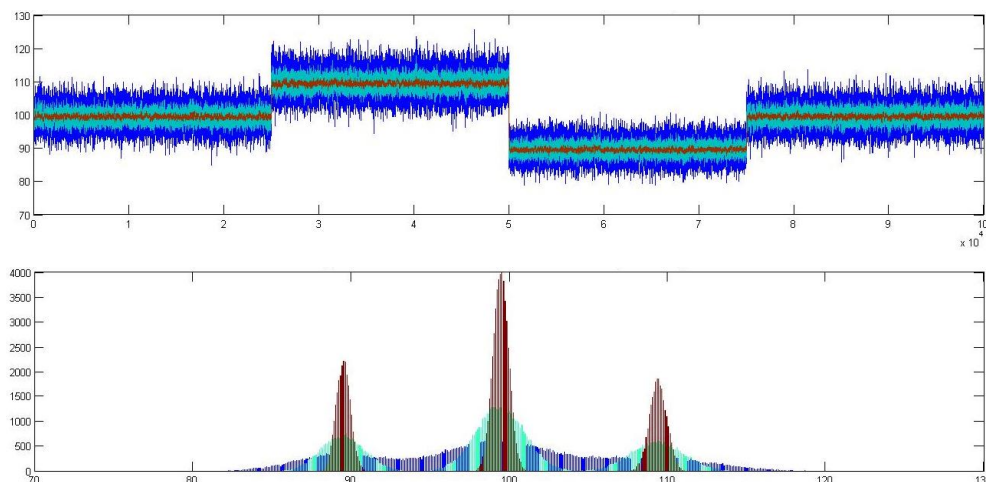
del evento o reducción en la respuesta de control ante la evidencia de una fuga, sea operativa o por acción de la delincuencia. Se simulará entonces cambios de flujo del 10% para pérdidas y ganancias. Estos cambios serán instantáneos para simular la condición más adversa de velocidad para el filtro. En condiciones normales, las fugas operativas o pérdidas por hurto son menores al 1% y son progresivas en el tiempo.

Tabla 4-3: Filtro Kalman para varios factores de varianza q a variaciones rápidas

q	0	0,1	0,05	0,01	0,005	0,001	0,0005	0,0001
Promedio	99,4211	99,4211	99,4211	99,421	99,421	99,4209	99,4209	99,421
Máximo	125,6053	121,1276	120,1183	117,7345	116,6707	114,0706	113,0903	111,4217
Mínimo	78,8215	80,566	81,3942	84,0128	84,9949	86,1632	86,6093	87,6613
Desv.Est	7,7674	7,4634	7,3674	7,1956	7,1483	7,0817	7,0651	7,0407
Varianza	60,332	55,7017	54,2786	51,7765	51,0977	50,1508	49,9155	49,572
Curtosis	2,3443	2,2201	2,1764	2,0925	2,0681	2,0325	2,0236	2,0115

Como se ha visto en las pruebas a flujo constante y bajas frecuencias, el filtro Kalman incide poco en las características de medición como el promedio. Dada las condiciones inducidas al sistema tanto como los límites y la desviación estándar no serán tomados para esta prueba como tal dado que existen mediciones muy diferentes para todos los momentos de flujos diferentes, sería necesario revisar cada segmento para analizar adecuadamente. En la figura 4.6 se mostrará la respuesta de la simulación para los cambios rápidos de flujo.

Figura 4-6: Filtro Kalman para mediciones de flujo Interlinea a variaciones rápidas³.



³ Azules la señal cruda, Turquesa con covarianza de 0.05 y Rojo con una covarianza de 0.0001.

4.2.2 Evaluación de Filtro Ponderado para mediciones Interlínea

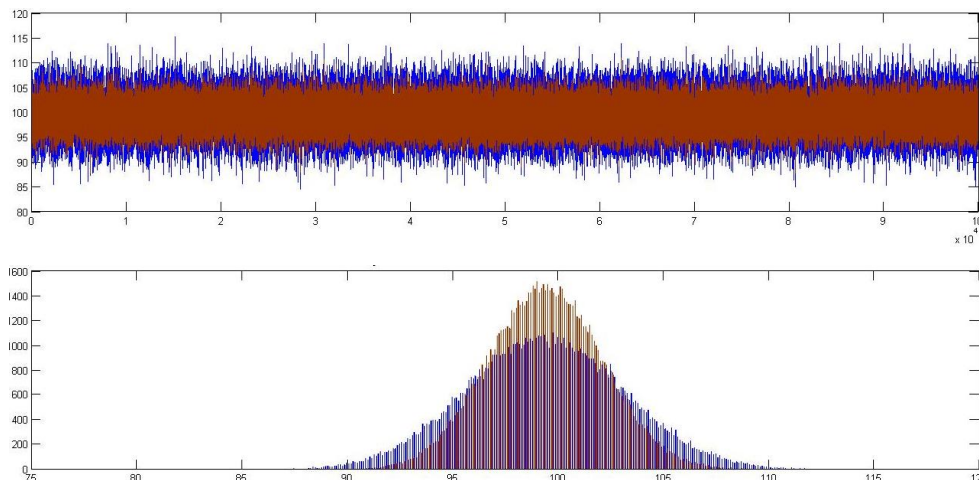
Para la evaluación de este filtro, se tomará la misma metodología que para el filtro Kalman. Este filtro al ser menos robusto, el único parámetro a variar será el factor A . Dada su simpleza, se evaluarán las tres condiciones; constante, cambios lentos y rápido en función de las variaciones porcentuales de A .

Tabla 4-4: Filtro Ponderado a varios factores de A .

A	1	0,2	0,3	0,4	0,5	0,6	0,7	0,8
Promedio	99,4243	99,4243	99,4243	99,4243	99,4243	99,4243	99,4243	99,4243
Máximo	115,5248	112,6449	111,9929	111,3408	111,1205	112,0014	112,8822	113,7631
Mínimo	84,9654	86,7148	87,2252	87,7356	88,246	88,1411	87,5483	86,9554
Desv.Est	3,7741	3,1151	2,8784	2,7265	2,674	2,7265	2,8784	3,1151
Varianza	14,2438	9,7038	8,2851	7,4338	7,1501	7,4339	8,2851	9,7038
Curtosis	2,9693	2,9771	2,985	2,9919	2,9926	2,986	2,9767	2,9702

Se puede ver en la respuesta del filtro no afecta el promedio acumulado de las muestra, lo que es óptimo para la reducción de ruido en mediciones de flujo. Por parte de la efectividad de este, cuando se aplica un factor $A = 0,5$ se analiza un máximo relativo en la reducción de la desviación estándar. Esto indica que este factor es el máximo de respuesta del filtro ante la variación de observaciones hechas por el medidor. En la figura 4.7 se mostrará la respuesta del filtro ante señales contantes de flujo.

Figura 4-7: Filtro Ponderado para mediciones de flujo⁴.



⁴ Azules la señal cruda y Rojo con factor $A = 0,5$.

Según la anterior gráfica y tablas de análisis, el uso de este filtro se ve estrictamente limitado al factor A en valores cercanos a 0,5. Esta limitante restringe el uso de este filtro para esta metodología por su bajo impacto en la reducción de incertidumbre de las muestras de apenas $0,3\sigma$ de la muestra

Filtro Ponderado para mediciones Interlinea con variaciones a bajas frecuencias

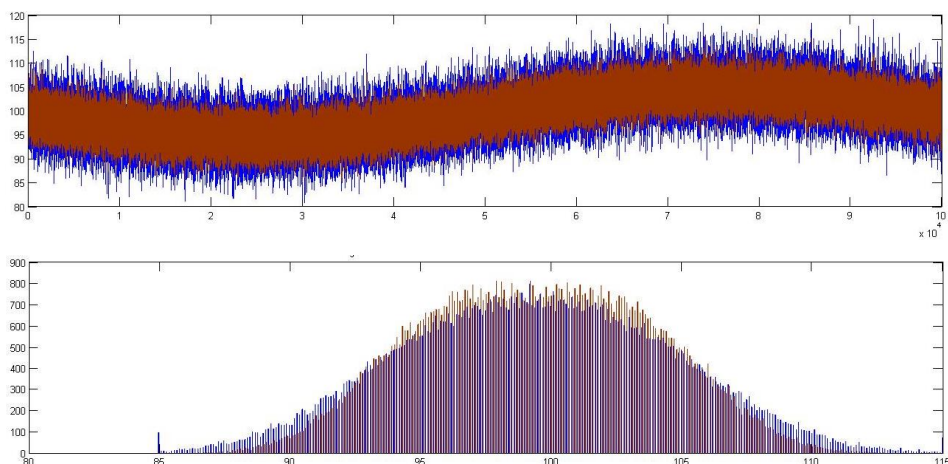
Tanto como se hizo con el filtro Kalman, se procederá a evaluar el filtro Ponderado en condiciones de cambio lentos de flujo en un DTHR. Esta evaluación nos permitirá entender como este filtro se adapta a cambios lentos de flujo por incidencias operacionales normales en un DTHR.

Tabla 4-5: Filtro Ponderado a varios factores de A para bajas frecuencias.

A	1	0,2	0,3	0,4	0,5	0,6	0,7	0,8
Promedio	99,4322	99,4322	99,4322	99,4322	99,4322	99,4322	99,4322	99,4322
Máximo	118,8456	117,333	116,5767	115,9081	115,6654	115,4227	115,9845	116,9382
Mínimo	79,171	82,6979	84,4614	84,5492	84,0581	83,4829	82,9076	82,3324
Desv.Est	5,1552	4,6932	4,5392	4,4442	4,4121	4,4442	4,5391	4,6932
Varianza	26,576	22,0259	20,6039	19,7507	19,4663	19,7506	20,6038	22,0258
Curtosis	2,6549	2,5094	2,4439	2,3978	2,381	2,3976	2,4435	2,509

Al descentralizar las medidas un 5% sobre una onda sinusoidal habrá una dispersión de las muestras sobre la misma e impactando la varianza y los límites de los datos. El impacto al promedio es nulo, lo cual evidencia un buen comportamiento del filtro en el compendio total de las muestras, y al ser los cambios lentos, se induce que lo hará también en muestras parciales y continuas de la misma. En el figura 4.8 se puede apreciar la respuesta del filtro.

Figura 4-8: Filtro Ponderado para mediciones de flujo a bajas frecuencias⁵.



El comportamiento descrito en el gráfico 4.8 describe una repuesta pobre del filtro a valor centrales de las muestra aunque se aprecia un mejora en los extremos de la misma. Este comportamiento sumado a la dispersión de la información observada conducen a que la respuesta del filtro no sea efectiva para los propositos de la metodología propuesta en esta tesis ya que los comportamientos variables del sistema son normales en operaciones normales de los DTHR.

Filtro Ponderado para mediciones Interlinea con variaciones a variaciones rápidas

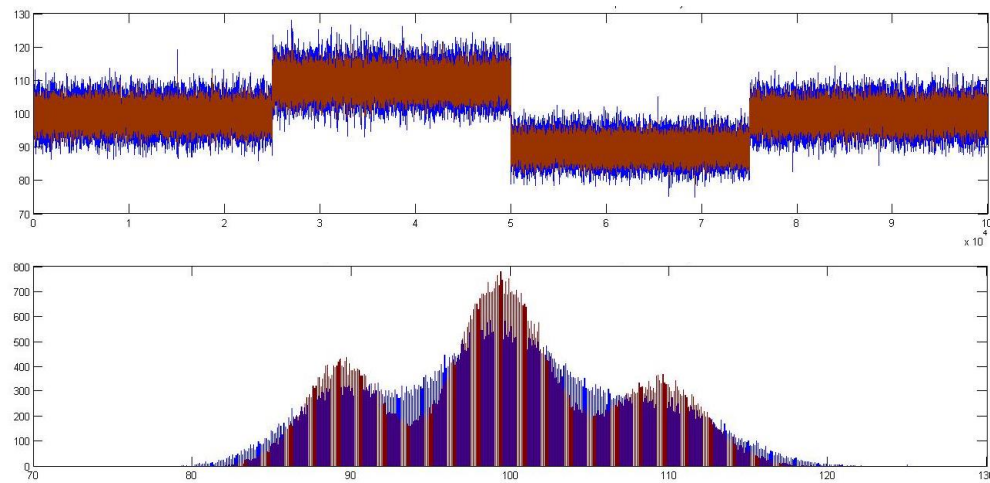
Tal como se describió la necesidad de esta prueba en el filtro Kalman, la velocidad de respuesta del filtro ante los cambios rápidos de flujo necesarios dada las condiciones externas inducidas por terceros a DTHR y siendo el fin de la método a aplicar en esta metodología.

Tabla 4-6: Filtro Ponderado a varios factores de A para variaciones rápidas.

A	1	0,2	0,3	0,4	0,5	0,6	0,7	0,8
Promedio	99,4215	99,4215	99,4215	99,4215	99,4215	99,4215	99,4215	99,4215
Máximo	127,9312	124,4431	123,9302	123,4174	122,9046	122,3917	122,2265	123,8982
Mínimo	77,1834	79,3408	80,1477	80,3262	79,9631	79,6001	79,1767	78,6107
Desv.Est	7,9728	7,6819	7,5888	7,5323	7,5134	7,5323	7,5888	7,6819
Varianza	63,5648	59,0119	57,5892	56,7355	56,451	56,7356	57,5892	59,012
Curtosis	2,414	2,31	2,2722	2,248	2,2395	2,2475	2,2714	2,3092

⁵ Azules la señal cruda y Rojo con factor A = 0,5.

Figura 4-9: Filtro Ponderado para mediciones de flujo a variaciones rápidas⁶.

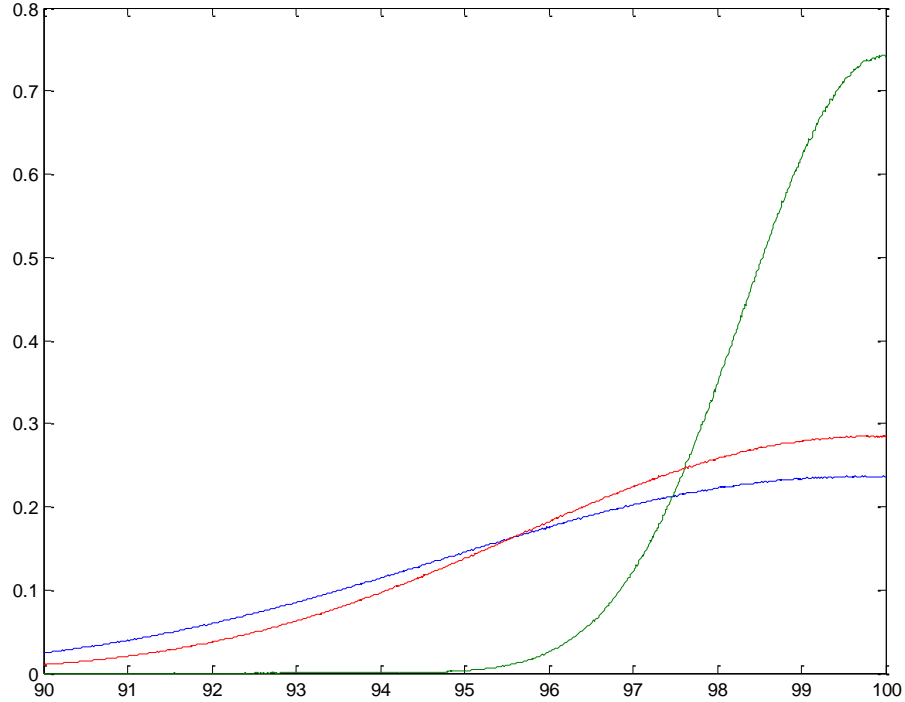


Se mantiene la tendencia en la cual el factor A de 0,5 es el que reduce mejor la dispersión de la información sin afectar claramente la medición en las muestras observadas en conjunto. Este filtro podrá utilizarse en fases posteriores a los análisis de información donde la capacidad computacional sea muy baja, por ejemplo en sistemas donde solo PLC de tercera generación sean los únicos idóneos para la tarea continua.

4.2.3 Filtros Kalman y Ponderado usando desigualdad de Chebyshev.

En la figura 4.10 Se puede observar como la probabilidad (azul) de que una medida observada por un medidor de desviación estándar del 5% este dentro de la función de distribución de probabilidad normal de un medidor del 0.5% de desviación estándar. Al aplicar un filtro Kalman (verde) con un $q = 0.001$ y un filtro de peso ponderado de $A = 0,5$ se puede apreciar una respuesta más rápida de la probabilidad a cambios de flujo de pequeño valor (menores al 1%) en el rango del 90% al 100% flujo.

⁶ Azules la señal cruda y Rojo con factor $A = 0,5$.

Figura 4-10: Respuesta de intersección entre medidores terminales e interlínea⁷.

Esto da como resultado, que un equipo de una incertidumbre alta puede detectar cambio de flujo menores a su incertidumbre utilizando filtros de información combinados con la probabilidad de intersección de valores utilizando la metodología de Chebyshev en función de la probabilidad de que una muestra de un medidor de alta incertidumbre filtrado se encuentre en la función de probabilidad de un medidor de baja incertidumbre.

4.3 Evaluación de identificación de L/G por medidores m_i (nivel 2)

Esta evaluación partirá del modelo de L/G que se generó para el análisis en la metodología. Este modelo se simplificará al comparar solamente las observaciones hechas del medidor interlínea m_i contra el comportamiento de las observaciones previas hechas por los medidores terminales m_{ini} y m_{fin} .

⁷ Utilizando filtros Kalman (verde) y peso ponderado (rojo) para la misma medida sin filtro (azul).

Se buscará con esta evaluación verificar el comportamiento de la probabilidad de cambio de datos de un medidor de alta incertidumbre (5%) en un momento k determinado contra las características estadísticas de un número de momentos k definidos de los medidores inicial y final de baja incertidumbre (0,5%). Se analizarán dos escenarios donde las L/G se encontrarán en sentido del flujo antes y después del medidor interlinea m_i para diferentes volúmenes de desviación que van desde el 1% al 0,1% del flujo pasante.

Para el análisis inicial se definirá una cantidad fija de $K = 1000$ datos para garantizar una adecuada caracterización estadística de la medición de estos medidores. Posteriormente se analizará la sensibilidad de la detección en función del factor K . Este análisis será de gran ayuda cuando se analice el comportamiento de los flujos en condiciones simuladas en el Capítulo 5 dado que esta información no contiene una cantidad de datos tan grande como para caracterizar estas condiciones.

4.3.1 Acondicionamiento de observaciones hechas por medidor m_i

La información suministrada por el medidor interlínea es de alta incertidumbre, se procederá a filtrar la señal para que esta pueda ser comparable con las medidas que se tendrán por los medidores terminales. Las medidas se filtrarán utilizando un filtro Kalman con una covarianza $q = 0.0001$ tal como se comprobó en la sección 4.2 para llevar muestras de información de flujo de alta incertidumbre a niveles comparables a observaciones hechas por medidores de baja incertidumbre. En la tabla 4.7 se mostrará las características estadísticas de las señales que se tendrán para la evaluación del nivel 2 de la metodología.

Tabla 4-7: Características estadísticas para medidores a evaluar en Identificación.

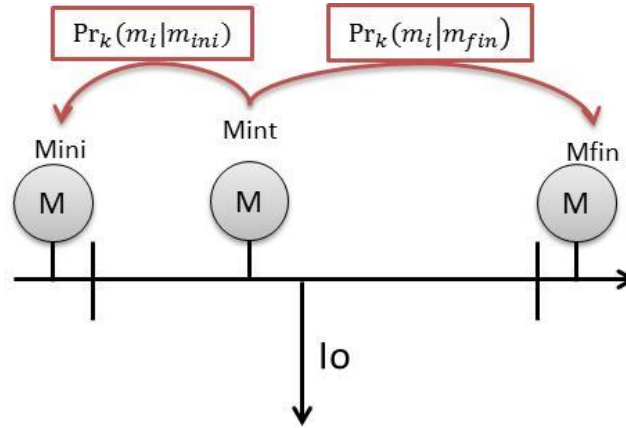
Medidor	Mini	Mfin	Mint	Mint_k
Promedio	100,0086	99,9913	99,9949	99,9926
Máximo	102,0452	101,8082	119,8434	102,6341
Mínimo	98,2249	97,9835	80,529	92,9196
Desv.Est	0,4925	0,4924	5,0686	0,7535
Varianza	0,2426	0,2425	25,6905	0,5678
Curtosis	2,9862	2,9998	3,0087	4,3625

Partiendo de la información de la tabla se puede evidenciar que los comportamientos de los medidores terminales son aproximados a los necesarios para esta evaluación. La información acerca del medidor interlínea (m_{int}) contrastada con la tomada después de filtrarla por medio de Kalman (m_{intk}), se puede evidenciar un desempeño estadístico óptimo para fines de esta evaluación dado que estos valores son ya cercanos a las características estadísticas de los medidores terminales. Se puede observar también una mejora considerable en la curtosis de la medición ya filtrada lo que indica un mayor apuntalamiento en la función de distribución que forman los valores. Como para la evaluación hecha para los filtros, se evidencia un cambio mínimo en la variación de esta característica estadística.

4.3.2 Evaluación de las mediciones de m_i cuando existe una L/G posterior

Se buscará en este análisis cómo una pérdida/ganancia provocada después del medidor interlínea m_i afecta los valores de probabilidad contrastados con las mediciones hechas por los medidores terminales. Para este análisis se generará una pérdida/ganancia de un 1% para evaluar el desempeño del uso de un filtro Kalman a los valores observados de un medidor interlínea. Posteriormente se analizará la sensibilidad en la variación de la probabilidad cuando las pérdidas/ganancias se reducen de 1% al 0,1%, este análisis determinará la sensibilidad teórica de la estrategia propuesta.

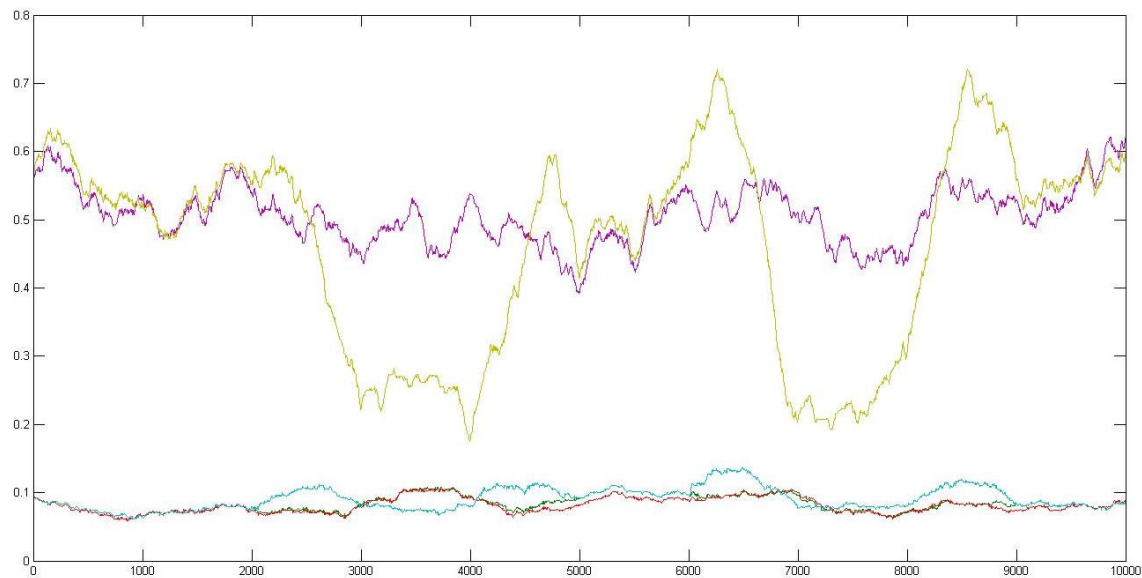
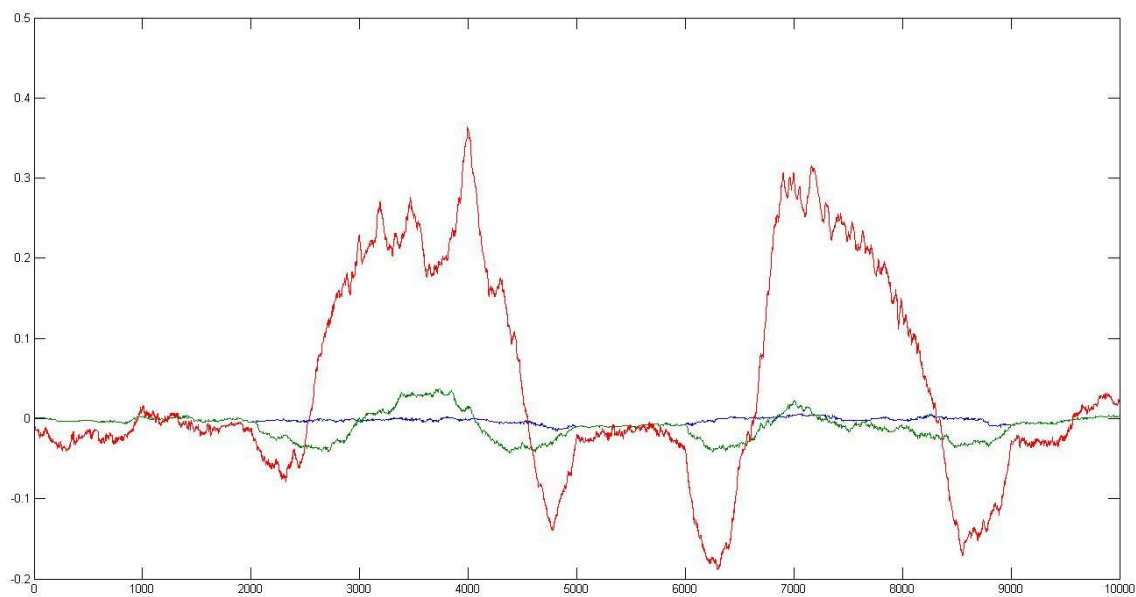
Generar una pérdida de un valor determinado posterior a la medición del medidor interlínea sugiere que las observaciones hechas por el medidor terminal final serán diferentes a los otros medidores del modelo simplificado. El valor de probabilidad variaría en función de la cantidad del valor de la L/G dado ya que habría un desplazamiento del valor real de las dos medidas.

Figura 4-11: Modelo simplificado para la medición de probabilidad de L/G posterior.

Se procede a generar valores con un flujo de 100 u/p y una pérdida/ganancia de un 1 u/p en diferentes momentos de la muestras en el segmento $m_{\text{int,fin}}$. Los medidores terminales cuentan con incertidumbres de 0,5% y el interlínea del 5%.

Se simulará un flujo que contiene 10000 momentos para el modelo propuesto en la figura 4.11; para verificar la sensibilidad del método propuesto se generarán una pérdida de 1 u/p entre el momento 2001 y el 4000 y una ganancia de 1 u/p entre el momento 6001 y el 8000. La sensibilidad será analizada partiendo de un modelo a.) sin L/G, b.) con L/G m_i sin filtro y c.) con L/G m_i con filtro Kalman con $q = 0.0001$. En la figura 4.12 se presenta el resultado de las probabilidades del medidor m_i contra las mediciones hechas por m_{ini} y m_{fin} teniendo en cuenta que la medición de m_{fin} tiene una diferencia de 1 u/p en sectores ya definidos.

En el figura 4.13 se presenta la diferencia aritmética para cada momento k entre las probabilidades $Pr_k(m_i|m_{\text{ini}})$ y $Pr_k(m_i|m_{\text{fin}})$ para el marco de pruebas propuesto anteriormente. Esta gráfica permitirá una mejor comprensión de la respuesta del modelo.

Figura 4-12: Valor de la probabilidad con y sin L/G posterior al medidor m_{int} ⁸.**Figura 4-13:** Diferencia de probabilidades de la figura 4.12⁹.

⁸ Valor de la probabilidad como si no se hubiese generado la nombrada desviación L/G (azul y verde) posterior al medidor m_{int} , el valor de la probabilidad del medidor interlínea sin filtro (rojo y cian) y con filtro Kalman (Amarillo y violeta).

⁹ Diferencia de probabilidades para cada momento k de la figura 12 para simulación sin pérdida/ganancia (azul) posterior al medidor m_{int} , con pérdida ganancia sin filtro (verde) y con pérdida ganancia con filtro Kalman (rojo).

En la figuras 4.12 y 4.13 se puede observar la respuesta del método aplicado en función de la aplicación del filtro Kalman realizado a las observaciones de m_i . Se puede analizar de estas figuras lo siguiente:

- El método propuesto no diferencia si existe una pérdida o ganancia para las desviaciones, la respuesta es homogénea para ambas eventualidades. Esto apalanca el objetivo de este nivel dado que solo se busca en esta la identificación de pérdidas/ganancias, la caracterización será llevada por el nivel 4.
- En las zonas donde no se encuentra L/G, las probabilidades varían poco con respecto a las zonas donde se encuentran las desviaciones. Esto significa que el modelo tiene sensibilidad a las L/G.
- Los valores filtrado tiene una mayor probabilidad contrastados contra las mediciones de baja incertidumbre que las no filtradas, esto garantizará que cualquier movimiento en el valor esperado produzca un mayor efecto en este indicador.
- Para las señales filtradas con Kalman se puede observar una mayor velocidad de respuesta que para las no filtradas cuando existe una L/G. Esta particularidad permitirá posteriormente una mejor identificación de L/G a menores tasas de desviación.
- Se observa que una vez existe una L/G, los valores de probabilidad varían inicialmente positivo y posteriormente varían negativamente después de unos momento k determinados. Este amortiguamiento de la señal inicialmente es debido la cantidad de datos K definidos para la simulación de 1000. Si se reduce este valor, los valores pico relativos se pueden reducir aunque se sacrificaría la confiabilidad del método dado que los valores K generan los límites de probabilidad del mismo análisis.
- Los efectos del amortiguamiento se presentan al inicio y final de los valores generados de L/G. Se podría establecer que la duración o periodo de una desviación L/G se establecería como:

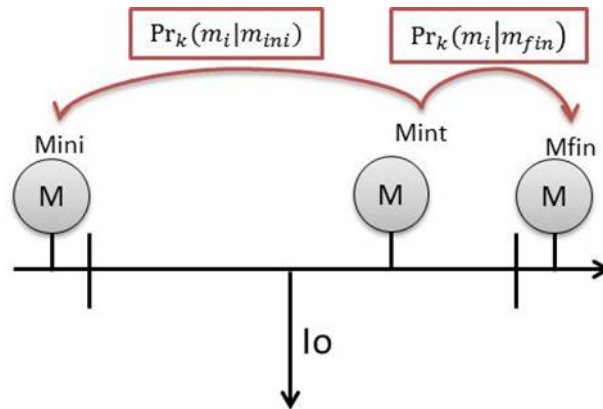
$$L/G_k = k_{fin} - k_{ini} - K \quad (4.1)$$

- Al analizar la figura 4.13, se puede observar un mejor comportamiento de las señales de probabilidad dado que los efectos aleatorios de la naturaleza de estas señales se compensan entre ellas mismas. En función de esta variable se podrá posteriormente generar las alarmas de identificación de pérdidas/ganancias.

4.3.3 Evaluación de las mediciones de m_i cuando existe una L/G anterior

Analizaremos ahora el efecto en las probabilidades $Pr_k(m_i|m_{ini})$ y $Pr_k(m_i|m_{fin})$ cuando la pérdida/ganancia es generada entre el medidor inicial y el medidor interlínea. Para plantear conceptualmente se generara el siguiente modelo de análisis (figura 4.14).

Figura 4-14: Modelo simplificado para la medición de probabilidad de L/G anterior.



Partiendo del mismo marco de evaluación propuesto en la sección 4.3.2, se simulará la L/G en el segmento que limita las mediciones del medidor inicial e interlínea. Para darle consistencia al análisis se mantendrá las otras características constantes. Los resultados de las simulaciones hechas se consolidarán en las figuras 4.15 y 4.16.

Figura 4-15: Valor de la probabilidad con y sin L/G posterior al medidor m_{int} ¹⁰.

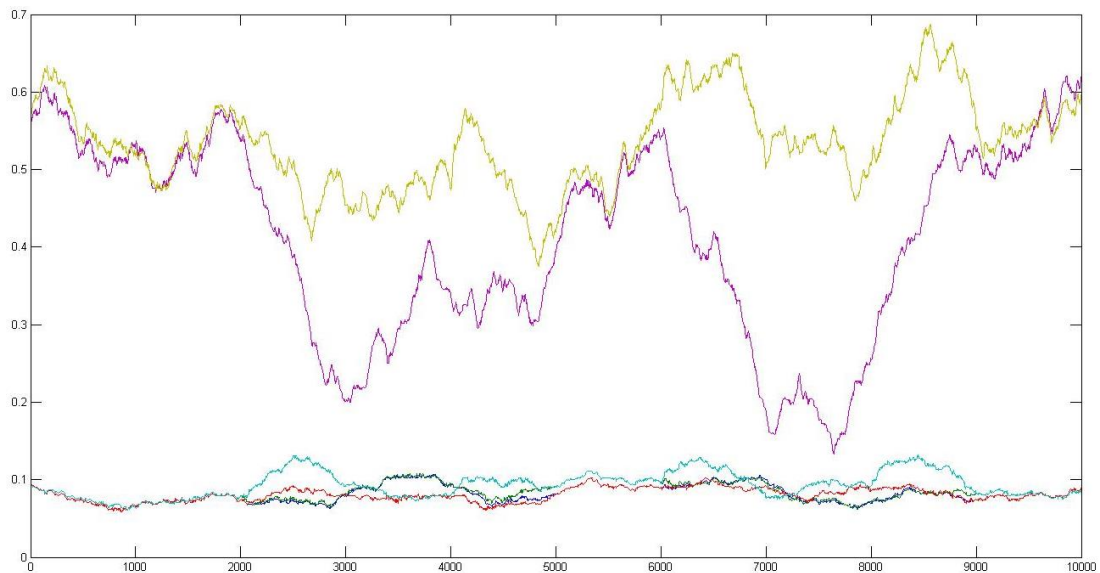
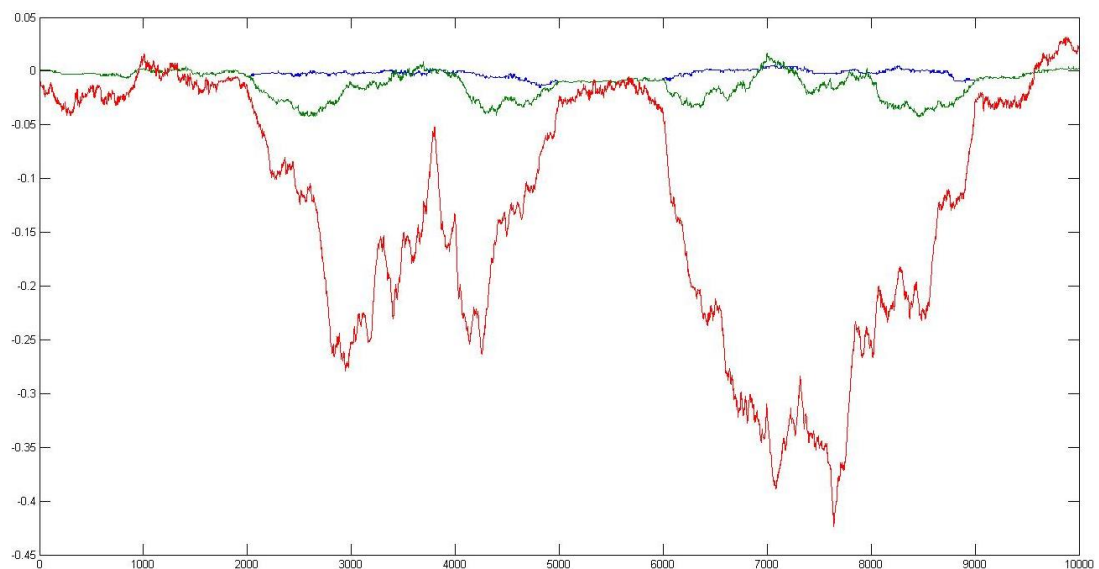


Figura 4-16: Diferencia de probabilidades de la figura 4.15¹¹.



¹⁰ Valor de la probabilidad como si no se hubiese generado la nombrada desviación L/G (azul y verde) anterior al medidor m_{int} , el valor de la probabilidad del medidor interlínea sin filtro (rojo y cian) y con filtro Kalman (Amarillo y violeta).

¹¹ Diferencia de probabilidades para cada momento k de la figura 4.15 para simulación sin pérdida/ganancia (azul) anterior al medidor m_{int} , con pérdida/ganancia sin filtro (verde) y con pérdida/ganancia con filtro Kalman (rojo).

Dado que la respuesta del modelo es homogénea a la descrita en la sección 4.3.2, el análisis solo se hará basado en características precisas del modelo. Podemos observar del resultado lo siguiente:

- Dado que la pérdida/ganancia se encuentra en el segmento anterior al medidor interlínea, los valores de probabilidad cambian para los valores $Pr_k(m_i|m_{ini})$ y $Pr_k(m_i|m_{fin})$. Esto es lógico en el sentido que los valores observados del medidor interlínea y final tiene un valor esperado igual y diferente al valor esperado de las observaciones hechas por el medidor inicial.
- Para el caso particular de la simulación, se encuentra una respuesta más marcada a los valores de ganancia que para los de pérdida. Esto respondería a la tendencia de las probabilidades una vez se genere el evento de pérdida/ganancia. Esta característica podría posteriormente afectar las mediciones de valores de L/G de menor magnitud para señales no filtradas. Por tanto el uso del filtro será necesario.

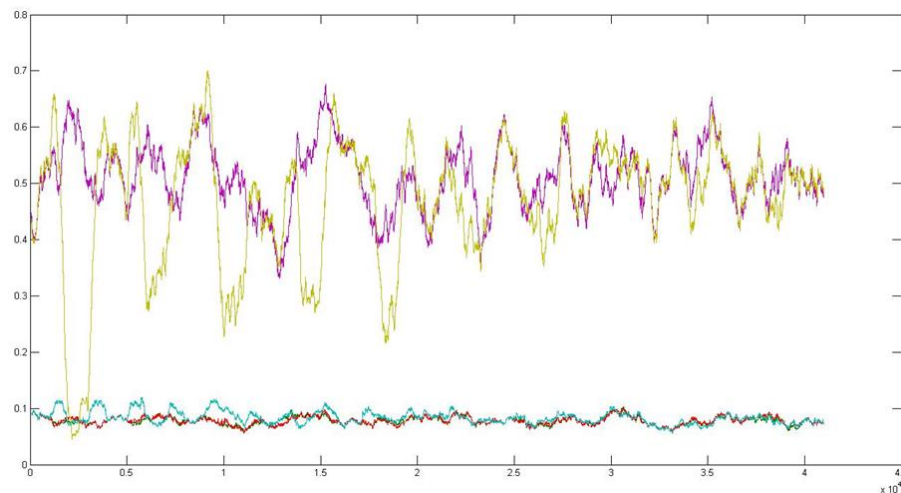
4.3.4 Evaluación de sensibilidad de identificación de pérdidas/ganancias

Se buscará con esta sencilla prueba generar una aproximación gráfica para interpretar la incidencia en la probabilidad cuando se generan pérdidas/ganancias de diferente magnitud sobre la línea. Esta prueba nos dará los insumos para generar estrategia para la supervisión de estado de los sistemas, inicialmente se partirá de un modelo de umbrales y bandas muertas para el control de estas variables. Otras estrategias podrían ser utilizadas posteriormente y estarían por fuera del alcance de esta tesis.

Para verificar la afectación del método de medición de probabilidad en función de la rata de pérdida o ganancia en un momento k determinado se realizará un análisis puntual para varias ratas. Este análisis se hará partiendo del modelo para pérdida posterior al medidor interlínea tal como lo muestra la figura 4.11. Se simulará un flujo nominal del 100% por 43000 momentos, se generará pérdidas de 2000 momento cada una por valores del 1% al 0,1% con diferenciales de 0,1% descendiendo. Se tomaron como variables de incertidumbre de 0,5% para los medidores terminales con un PDF cercano a

una distribución normal y 5% para el medidor interlínea con un PDF multivariable. Los resultados gráficos pueden ser vistos a continuación (figura 4.17).

Figura 4-17: Valor de la probabilidad para pruebas de sensibilidad¹².



Partiendo de la respuesta gráfica de la simulación de sensibilidad, se pueden hacer las siguientes observaciones:

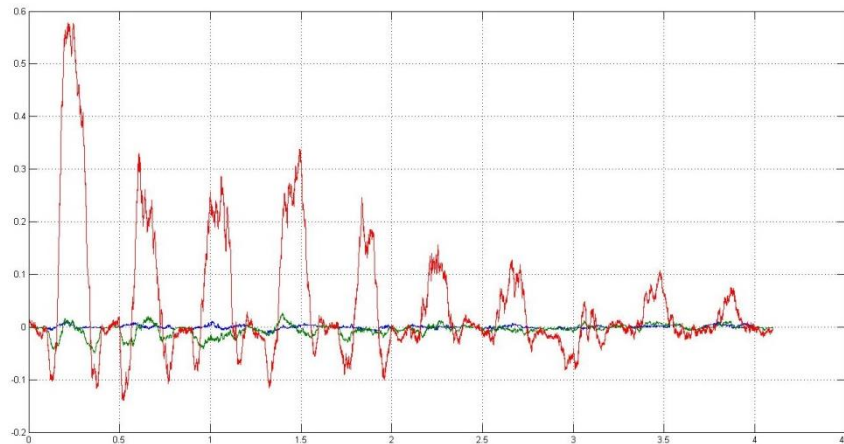
- El método propuesto es sensible ante el cambio de rata de pérdida/ganancia generado. Se verifica que mientras para una rata de flujo de 1% la respuesta es clara para las observaciones filtradas (amarillo) y no filtrada (cian) del medidor interlínea son evidentes con respecto a la señal si un hubiese pérdida (verde), para una pérdida/ganancia de 0,1% son casi imperceptibles.
- A cambios menores de las ratas de pérdida/ganancia, tanto para las observaciones no filtradas como las filtradas mismas, se evidencias afectaciones en la probabilidad de intersección de las observaciones por el método propuesto. Estos cambios difieren en magnitud a la rata de pérdida/ganancia; para señales de menor rata es imperativo el uso del filtro para magnificar la desviación presente.

¹² Valor de la probabilidad para pruebas de sensibilidad como si no se hubiese generado L/G (azul y verde) posterior al medidor m_{int} , el valor de la probabilidad del medidor interlínea sin filtro (rojo y cian) y con filtro Kalman (Amarillo y violeta).

- Se mantiene el efecto de amortiguamiento anterior y posterior una vez existe una pérdida, este se reduce a medida que se reduce la rata de la pérdida/ganancia. Se puede aseverar que este efecto proporcional depende directamente del factor K cuando se analiza los umbrales de probabilidad del valor en momento k . Este valor se podría dejar fijo dependiendo de los valores históricos de cada medidor pero se perdería adaptabilidad del modelo aplicado ante eventos propios del sistema. Se puede considerar este análisis para pasos posteriores fuera del alcance de esta tesis.
- El comportamiento de las probabilidades cuando no existe pérdida se limita en umbrales definidos, entre un 5% y 11% para señales no filtradas, se puede establecer estos límites como puntos de control para generar alarmas de estado del sistema de medición en particular.

Para una interpretación más consistente, se mostrará en la gráfica a continuación (figura 4.18) la sustracción aritmética de los valores de probabilidad $Pr_k(m_i|m_{ini})$ y $Pr_k(m_i|m_{fin})$. Esta operación desligaría el efecto aleatorio en cierta proporción pero no el efecto aleatorio inducido por el mismo filtro. Se puede observar con mayor claridad la respuesta del método propuesto a ratas de pérdidas/ganancias de menor magnitud.

Figura 4-18: Diferencia de la probabilidad para pruebas de sensibilidad¹³.



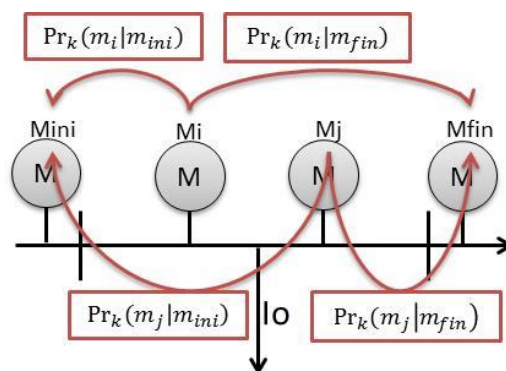
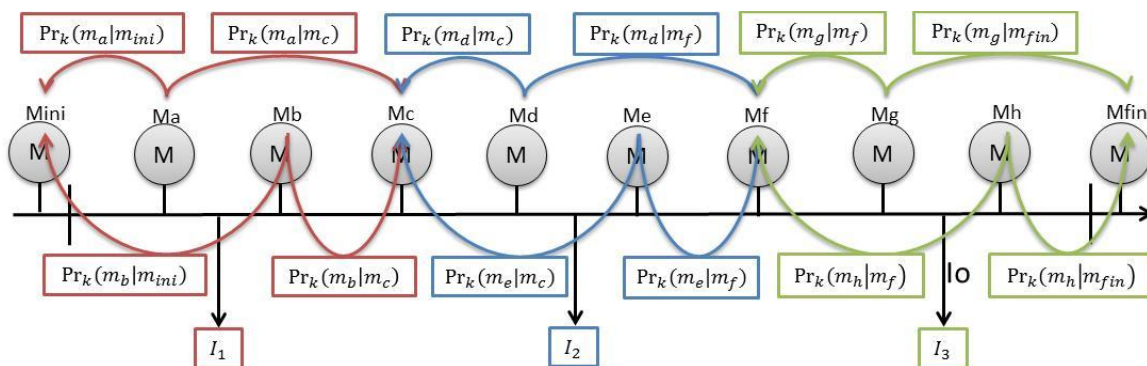
¹³ Diferencia de probabilidades para análisis de sensibilidad para simulación sin pérdida/ganancia (azul) posterior al medidor m_{int} , con pérdida/ganancia sin filtro (verde) y con pérdida/ganancia con filtro Kalman (rojo)

4.4 Evaluación de análisis de aislamiento de L/G para segmentos m_{ij} (nivel 3)

Dado que la una falla de un medidor en particular, pérdida de calibración o falla general, puede generar la evidencia de una pérdida/ganancia ficticia, es necesario el análisis del comportamiento de este medidor contra sus pares, sean medidores terminales o interlínea adyacentes al mismo. Los análisis hechos para comprobar la potencialidad de este método propuesto van en función de descartar fallas propias de los medidores interlínea basada en la fusión de la información de los medidores adyacentes contrastados contra los medidores terminales en casos de una sola falla o adyacente para múltiples fallas en el DTHR.

Para mantener la homogeneidad con las pruebas hechas para los anteriores niveles, se mantendrán los niveles de incertidumbre de 0,5% para los medidores terminales y 5% para los medidores interlínea. Estas definiciones permitirán llevar una mejor comprensión de los resultados presentados al nivel de análisis de resultados pero pruebas a diferentes niveles de incertidumbre fueron realizadas fuera del marco definido.

Tal como se estipuló anteriormente, se tendrá dos posibles eventualidades para la fusión de la información de los diferentes segmentos definidos por los medidores. Las L/G pueden ser únicas o múltiples para un momento determinado, el aislamiento de estas desviaciones deben tener un tratamiento diferente debido a la mediciones hechas por el medidor de referencia, terminales para L/G única adyacentes para múltiples L/G. En la figura 4.19 se podrá esclarecer la forma como se hará la fusión de información partiendo de los medidores de referencia para una sola desviación L/G. En la figura 4.20 se podrá observar el tratamiento de los puntos de referencia cuando exista más de una desviación L/G, en este caso y para 8 medidores interlínea, tres desviaciones de pérdida en el mismo momento.

Figura 4-19: Método de fusión de información partiendo de medidores terminales.**Figura 4-20:** Método de fusión de información partiendo de medidores Interlínea.

Para todos los casos en que se utilice observaciones hechas por los medidores interlínea se debe hacer uso de filtros Kalman para garantizar la sensibilidad ante bajas ratas de desviaciones L/G. El uso de este filtro fue comparado en el nivel 2 de la metodología y por tanto se da la viabilidad para el uso de este.

La evaluación del método propuesto será llevada en dos instancias. Inicialmente se analizará la efectividad y sensibilidad del método para el modelo propuesto en la gráfica 4.19 de una única desviación para los métodos de fusión de información por Bayes y Dempster-Shafer. Posteriormente se analizará la confiabilidad del aislamiento de desviaciones basados por en el modelo propuesto en la gráfica 4.20 por el método de fusión que haya presentado un mejor desempeño en el análisis anterior. Para este análisis final se harán diferentes pruebas a ratas de ganancias y desviaciones diferentes para comprobar su efectividad en el aislamiento de las mismas.

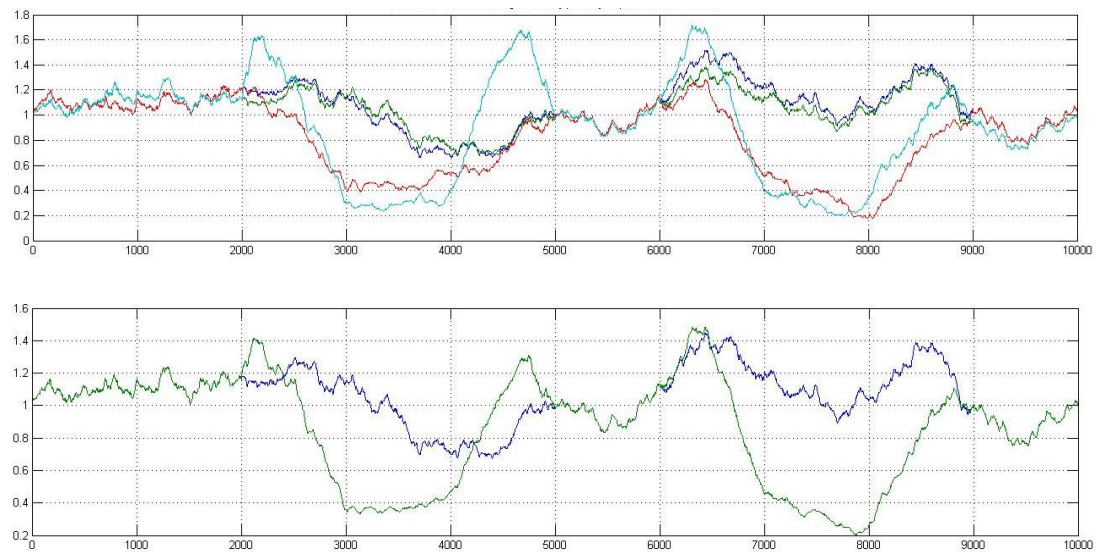
Finalmente se verificará la efectividad de aislamiento de ambos métodos de fusión en una simulación de una línea con diferentes segmentos. Esta prueba mostrará el estado de L/G de cada segmento, lo cual permitirá discernir el comportamiento contrastado de cada segmento contra sus adyacentes.

4.4.1 Evaluación de aislamiento de L/G en segmentos m_{ij} por BY

El método de fusión de valores por Bayes permite la fusión directa de las probabilidades de cada medidor en su primera desviación estándar. Dado que el intervalo de confianza de una desviación estándar para una distribución de probabilidad aproximada a una normal es mayor al 65%, solo se tomará este intervalo para generar la fusión.

Se simulará entonces una pérdida y ganancia de un 1% en diferentes momentos de para el modelo plasmado en la figura 4.19. Esto mostrará la efectividad del método en la detección de L/G en momentos diferentes con diferentes naturalezas. Los resultados se plasmarán en la figura 4.21.

Figura 4-21: Fusión de datos con referencia terminales con L/G 1% por BY¹⁴.

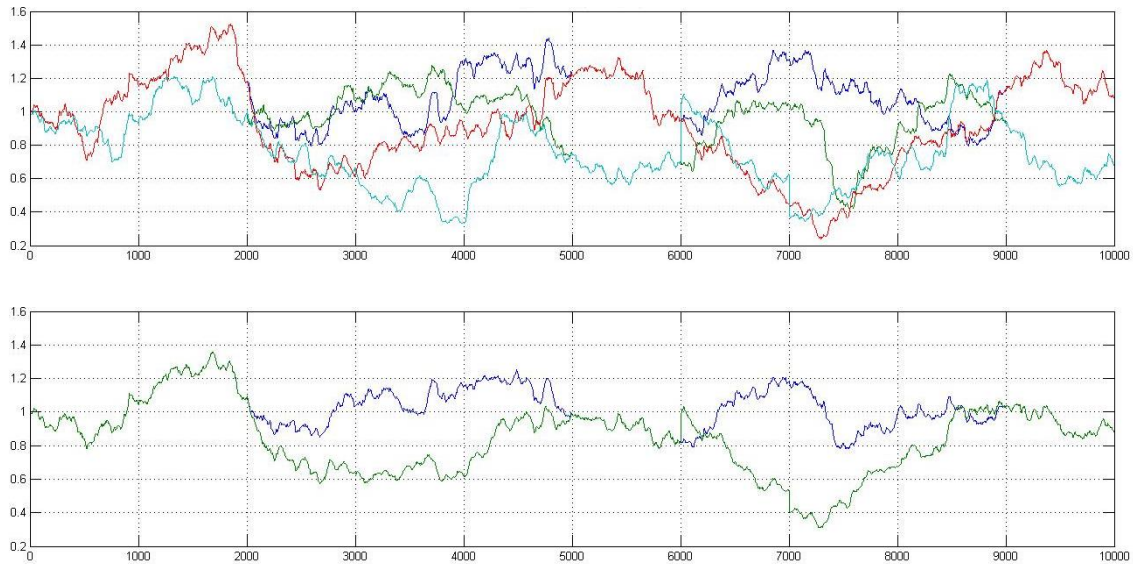


¹⁴ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Bayes para una L/G del 1% con medidores de referencia terminales.

Según los resultados dados por la fusión de la información por Bayes del segmento m_{ij} se evidencia lo siguiente:

- La fusión de la información tiene variaciones cercanas a $\pm 40\%$ (ver líneas azul y verde de la tendencia superior), esta variación podría esconder el comportamiento de L/G determinado para el segmento en la zonas mayores a uno.
- Se observa una caída de la probabilidad menor del 40% cuando existe la pérdida/ganancia, esto garantiza la aplicación del método Bayes para la fusión de datos observados de medidores en segmentos m_{ij} cuando existen L/G de 1% en estos.
- Cuando no existen comportamientos de L/G, las fusiones se comportan homogéneamente con variaciones menores entre ellas. Cuando existen L/G los comportamientos aunque homogéneos tienen variaciones muy altas. Este comportamiento evidenciaría una posible L/G en el segmento además de umbral de confianza descrito en el punto anterior.
- La fusión de valores lleva el intervalo de confianza de un medidor simple del 65% a valores mayores que 1%. Esto no tiene sentido lógico dado que la probabilidad no puede ser mayor al 100%.

Para el caso que los medidores de referencia no sean los medidores terminales sino uno de los interlínea las respuestas pueden cambiar para la fusión por el método Bayes. Para este caso se filtran las observaciones hechas por los medidores interlínea con filtro Kalman con una covarianza de $q = 0,0001$. Los resultados de esta prueba son presentados en la figura 4.22.

Figura 4-22: Fusión de datos con referencia interlínea con L/G 1% por BY¹⁵.

Aunque los resultados no son tan notorios para cuando se toma como referencia las mediciones de los medidores terminales, se encuentra un comportamiento homogéneo con referencia a un medidor interlínea con filtro Kalman. Esto implicaría una menor sensibilidad a las ratas de pérdidas/ganancia, lo cual conllevará a un factor determinante a la aplicación de esta estrategia para la fusión de datos en segmentos m_{ij} .

Sensibilidad de L/G para fusión Bayes para el segmento m_{ij}

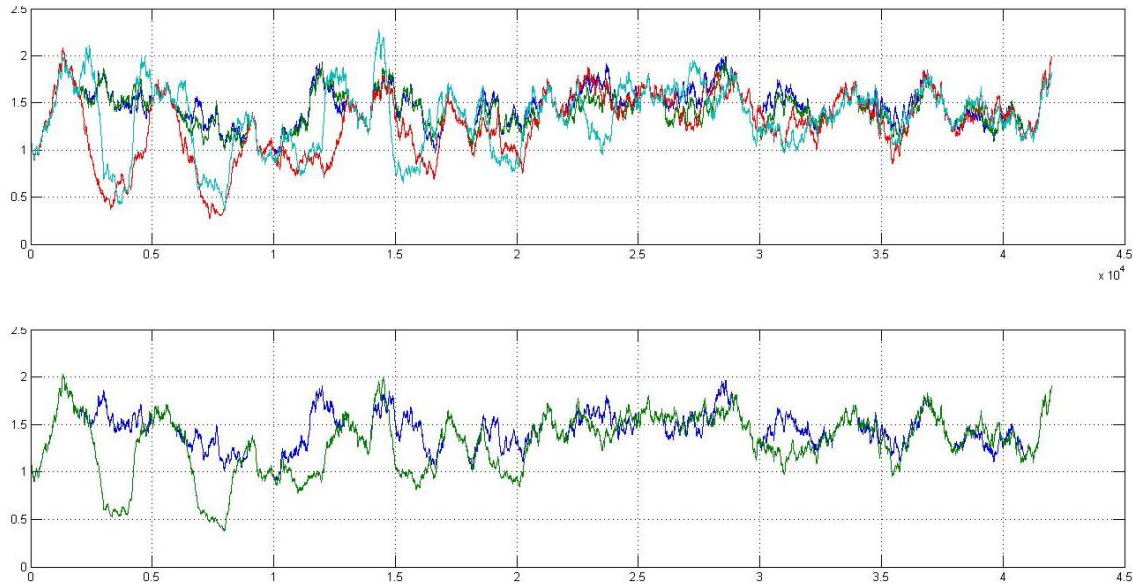
Con el fin de comprobar la sensibilidad de la fusión Bayes para las diferentes ratas de flujo de L/G se tomará el modelo de la gráfica 4.18 y se le simulará ratas de pérdida del 1% al 0.1% descendiendo en un ritmo de 0.1%. Estos nos permitirá visualizar hasta qué punto es útil utilizar un umbral de estado para determinar si hay una pérdida real o es solo una variación propia del efecto estadístico que se está generando.

Esta prueba considerará dos marcos de referencia diferentes, primero como medidor de referencia los terminales y el segundo con medidores interlínea. Estas dos simulaciones darán como resultado la sensibilidad cuando se tenga una sola pérdida y el otro cuando

¹⁵ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Bayes para una L/G del 1% con medidores de referencia interlínea.

existan múltiples pérdidas a lo largo del DTHR como se consignó anteriormente en función de las ratas de pérdida. Los resultados para el marco donde se encuentra tomando como referencia el medidor terminal se muestran en la figura 4.23.

Figura 4-23: Fusión para un segmento m_{ij} por BY con L/G del 1% al 0,1% T¹⁶.



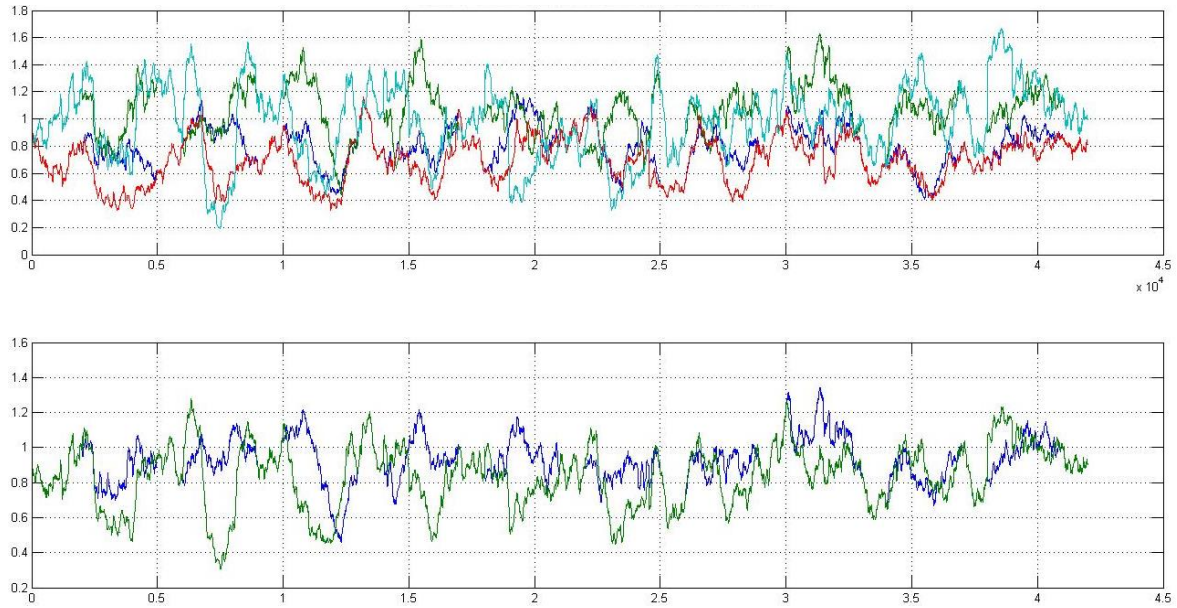
De los resultados después de aplicar la fusión de información por Bayes se pueden hacer las siguientes observaciones:

- Tomando el comportamiento del sistema si no existiera pérdida (línea azul gráfica inferior) y contrastándola contra la línea de aislamiento de pérdidas (línea verde gráfica inferior, los niveles de pérdida que van del 1% al 0,5% sobrepasan los máximos históricos de la línea que no tuviera pérdida, es decir, bajo una modalidad de umbral de estado se podría detectar estas desviaciones bajo el modelo de fusión de Bayes.
- Después de los niveles de pérdida de 0,5% en adelante no se puede discernir claramente si es un evento de este tipo bajo modalidad de umbral de estado. Sería necesaria otra estrategia como análisis de gradientes o comparadores de estado óptimo para detectar estas desviaciones.

¹⁶ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Bayes para varias L/G del 1% al 0,1% con medidores de referencia terminales.

Dada la simpleza computacional del método de fusión de Bayes este puede ser práctico para sistemas que no cuenten con mayor capacidad de la misma. Esta opción se vería limitada al aislamiento de eventos de L/G hasta un 0,5%.

Figura 4-24: Fusión para un segmento m_{ij} por BY con L/G del 1% al 0,1% ¹⁷.



Los resultados cuando se toma como referencia los medidores interlínea se presentan en la gráfica 4.24. Se puede verificar que la sensibilidad es menor cuando se utilizan medidores interlínea como referencia. Esto es debido al rendimiento del filtro Kalman que no llega hasta los niveles de precisión de un medidor terminal y por ende menos sensible a los cambios por desviaciones L/G. La fusión de información por Bayes no es adecuada cuando se utilizan medidores interlínea como referencia para L/G menores de 1%.

Aislamiento de única L/G por fusión Bayes para múltiples segmentos m_{ij}

Teniendo como punto de partida que las desviaciones L/G únicas se presentan con mayor frecuencia que las múltiples, se analizará esta condición de una manera más precisa y ampliada que para la condición múltiple. Para el diseño de esta prueba se tomará como punto de partida la metodología presentada en la gráfica 4.20 con la diferencia que se tendrán 6 medidores interlínea y los medidores de referencia serán los medidores

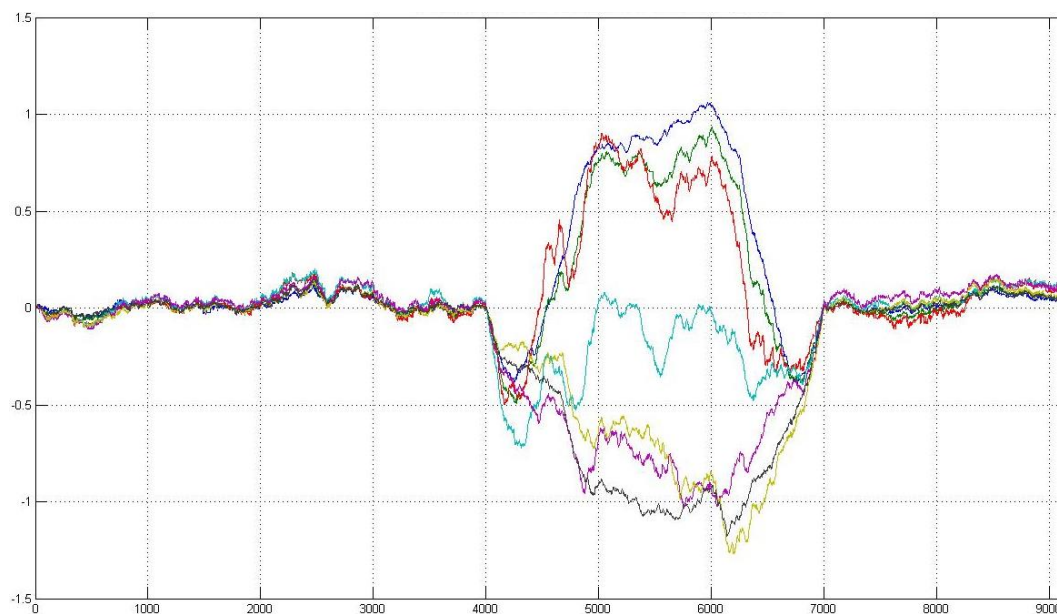
¹⁷ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Bayes para varias L/G del 1% al 0,1% con medidores de referencia interlínea.

terminales. En una segunda prueba se mostrarán los resultados si se tuviesen medidores interlínea como medidores de referencia.

Partiendo del uso de la fusión de información por Bayes, se tendrá un indicador preciso de cada segmento de un DTHR que será extraído de las probabilidades de los valores observados por los medidores que limitan el segmento contra los medidores terminales. Este indicador será presentado para cada segmento incluyendo los segmentos 1&2 y 7&8 que son los que se encuentran referenciados por ellos mismos.

Para incrementar la respuesta de los resultados, se realiza una sustracción aritmética entre las probabilidades referenciadas contra los medidores terminales. Esto además de llevar centrar los valores al origen permite reducir el ruido generado por la naturaleza aleatoria de las observaciones hechas por los medidores. En la gráfica 4.25 se presentan los resultados de la simulación donde se generó una pérdida de 1% por 2000 momentos desde el momento 4000.

Figura 4-25: Fusión datos múltiples segmentos m_{ij} por BY, única L/G de 1% T.

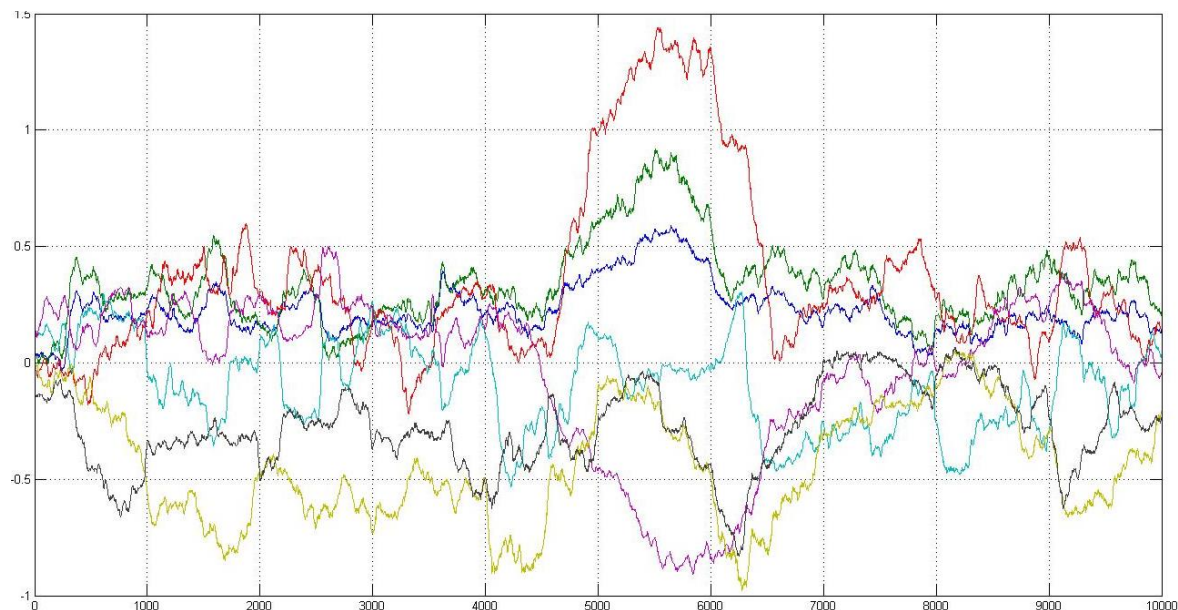


De los resultados presentados en la figura 4.25 se puede determinar lo siguiente:

- Dado que las tendencias provienen de resultados de las probabilidades restadas, las variaciones en los segmentos donde no hay desviación se encuentran muy próximas y centradas a cero.
- La línea del sector azul claro (data 4) son las observaciones fusionadas del segmento 4&5, donde se presentó la desviación. Se puede observar que esta tendencia es la de menor variación debido a que en este punto los valores de probabilidad deben comportarse idénticamente, mientras para los demás segmentos las probabilidades varían masa dado las diferencias de medición. Entonces el método de aislamiento de un segmento con desviaciones se puede determinar analizando la variación de los indicadores fusionados.
- La detección de una desviación es determinada por todos los segmentos del sistema, es decir, todos los medidores detectan esta desviación y son afectados por las mismas mediciones. Esto permite decir que el modelo es robusto en caso de fallas propias de medidores individuales dado que se puede anular la información de este y ampliar el sector. Esta práctica puede ser analizada posteriormente dado que esta por fuera del alcance de esta tesis.
- El factor K es determinante para definir el periodo de donde se presentó la desviación L/G. Para la prueba hecha, se definió como cantidad de momentos de 4000 momentos sin L/G. El periodo de inestabilidad de las muestras esta dado en 3000 momentos (del momento 4000 al 7000). Esto es debido que después de 1000 momentos los valores de referencia se vuelven a adaptar a las condiciones normales y se reorganizan nuevamente.
- Para definir el segmento que presenta la desviación en un periodo de inestabilidad R iniciando en un momento r_0 y finalizando en el momento r_n , se tomará el valor mínimo de los resultados para cada segmento m_{ij} de la ecuación 4.2 como el segmento que presenta la desviación L/G:

$$\min \left(\Delta M_{ij} = \sum_{r_0}^{r_n-K} \left| \widehat{m}_{ij_{ini}} - \widehat{m}_{ij_{fin}} \right| \right) \quad (4.2)$$

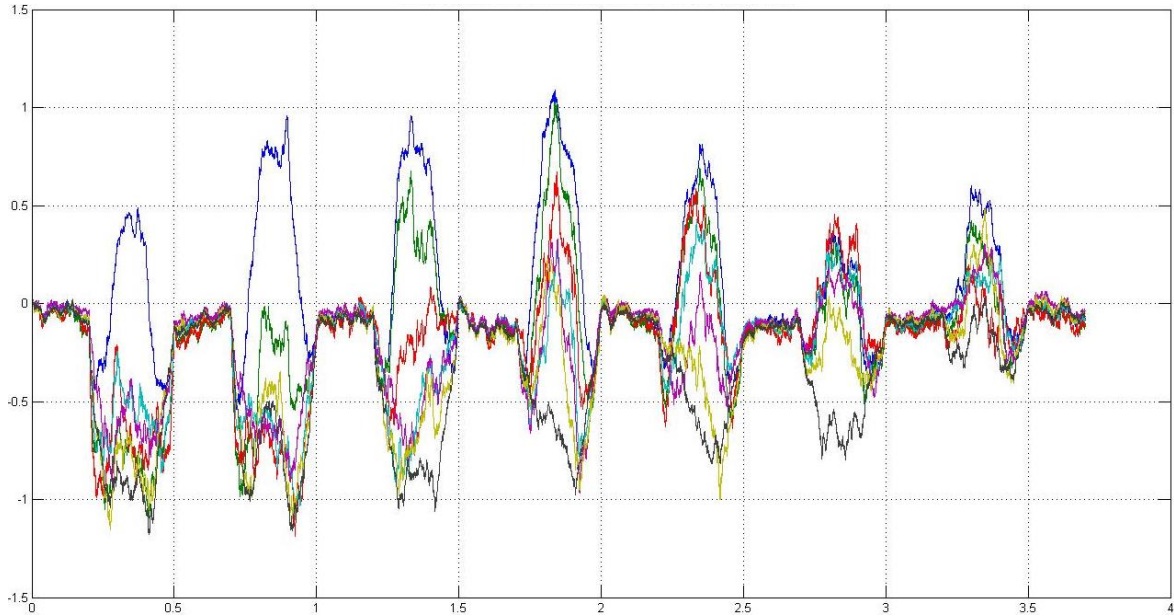
Para el caso en que los medidores de referencia sean los medidores adyacentes al segmento m_{ij} , es decir m_{i-1} y m_{j+1} , las respuestas dependen directamente del filtro que se utilice para refinar las observaciones hechas por los medidores interlínea prioritariamente. En la figura 4.26 se presentan los resultados de dicho ejercicio.

Figura 4-26: Fusión datos múltiples segmentos m_{ij} por BY, única L/G de 1% I.

Para estos resultados se puede resaltar que es complejo el análisis para aislar el segmento con desviaciones basadas en la observación únicamente. Utilizando la estrategia planteada para la ecuación 4.2 se tendría un resultado para el segmento 4&5 entre los momentos 4000 y 6000 de 307,9 mientras que para los otro segmento se tiene valores mayores a 600 hasta 1500.

Para el caso de múltiples L/G en diferentes momentos, se procederá con la simulación de diferentes L/G en todos los sectores del DTHR en análisis. Las L/G partirán desde un -1% en el sector 1&2 hasta un -0,4% en el sector 7&8. La información a simular de los medidores tendrá como medidor de referencia los medidores terminales.

Se busca con este análisis presentar diferentes niveles de pérdida en diferentes sectores para la fusión Bayes, que también serán replicados para Dempster-Shafer posteriormente, compactando varios resultados en una misma simulación; sensibilidad, aislamiento múltiple y fusión. Los resultados se presentan en la figura 4.27. Para una mejor estimación del análisis de la gráfica citada, se utilizará la ecuación 4.2 en la información fusionada y se agrupara estos resultados en la tabla 4.8 para posteriormente ser analizados.

Figura 4-27: Fusión datos múltiples segmentos m_{ij} por BY, múltiples L/G de 1% T.**Tabla 4-8:** Tabulación de datos de figura 4.27 según ecuación 4.2¹⁸.

Segmento(s) con L/G	1&2	2&3	3&4	4&5	5&6	6&7	7&8
Tamaño de L/G	-1,0%	-0,9%	-0,8%	-0,7%	-0,6%	-0,5%	-0,4%
P(INI,2,INI)-P(INI,2,FIN)	0,6432	1,1179	1,1514	1,2391	0,8757	0,3970	0,5937
P(2,3,INI)-P(2,3,FIN)	1,5300	0,8440	0,6904	0,9318	0,7359	0,3126	0,3601
P(3,4,INI)-P(3,4,FIN)	1,3780	1,4215	0,5422	0,5360	0,6484	0,4724	0,2313
P(4,5,INI)-P(4,5,FIN)	1,0887	1,2056	1,1291	0,4669	0,4917	0,3482	0,1915
P(5,6,INI)-P(5,6,FIN)	1,1268	1,1335	1,0067	0,5758	0,4518	0,2240	0,2190
P(6,7,INI)-P(6,7,FIN)	1,5555	1,2164	1,2783	0,5594	0,6320	0,4192	0,3618
P(7,FIN,INI)-P(7,FIN,FIN)	1,6829	1,3712	1,5118	1,0952	0,9526	1,1457	0,3352

Partiendo de la figura 4.27 y los resultados de las probabilidades fusionadas y restadas plasmadas en la tabla 4.8, se puede determinar lo siguiente:

- A medida que se reduce la rata de L/G se observa una afectación en la magnitud de la fusión de la información comparada, estas magnitudes son evidenciables hasta un valor de L/G de aproximadamente -0,5%. Aunque en la gráfica es visible la L/G en dicho segmento, ya al aplicar la ecuación 4.2 no es aplicable.

¹⁸ Resultados tabulados de fusión de datos de medidores de flujo para múltiples segmentos m_{ij} por fusión de Bayes para múltiples L/G de 1% en diferentes sectores con medidores de referencia terminales utilizando la ecuación 4.2.

- El esquema de aislamiento utilizando en la ecuación 4.2, se detectó L/G hasta en los segmentos 4 y 5. Esto indica que esta estrategia tiene límites de operatividad enmarcados básicamente por la aleatoriedad de las observaciones hechas por los medidores instalados en un DTHR.
- Se reconfirma lo dicho en el análisis de la gráfica 4.25, en donde las fusiones de las observaciones que varían poco en un periodo de L/G son las que presentan dicho L/G sin importar la posición en que se encuentren en un DTHR. Esto se verifica variando la posición de la L/G como se muestra en la tabla 4.8.
- Se observa que las probabilidades fusionadas por Bayes sobrepasan el valor de 1. Esto se evidenció en anteriores simulaciones y dado que el uso que se le está dando en no es para cuantificación podremos tomarlos como simplemente un indicador adimensional de estados de fusión de observaciones hechas por medidores que limitan un segmento.
- Aunque el método propuesto tiene una limitante para aislar L/G en el segmento en función de la sensibilidad ante rata de dicha desviación, es óptimo para identificar pérdidas a muy baja rata ($<0,4\%$), es decir, optimiza la detección usando varios medidores. Esta característica ofrece una alta potencialidad para detectar pérdidas en segmentos m_{ij} de un DTHR.

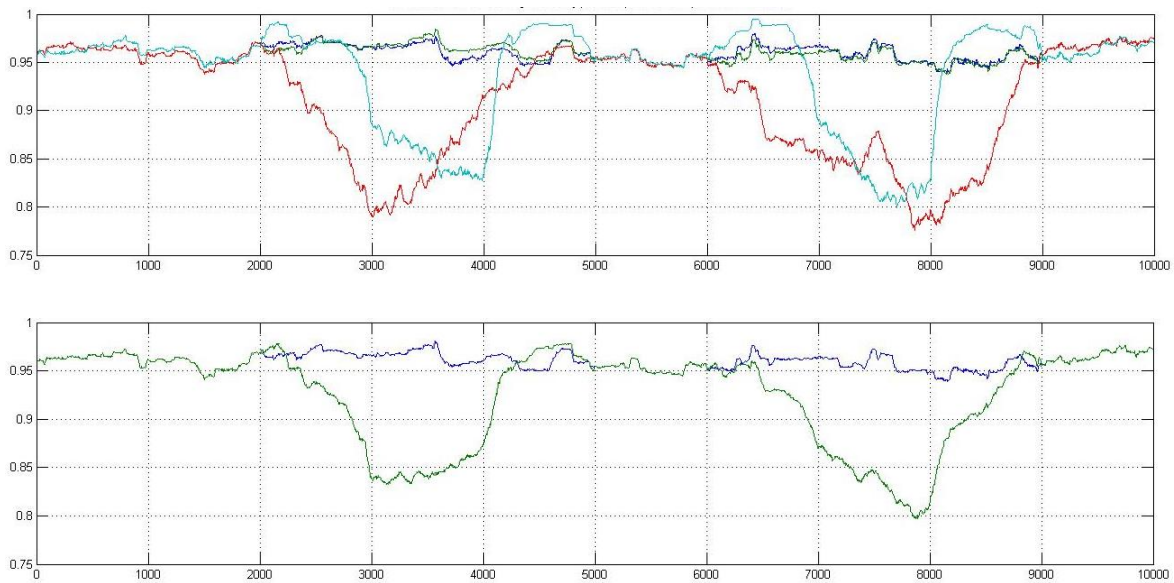
Para el caso particular y poco frecuente en operación de un DTHR de pérdidas simultáneas no se probará dada la poca probabilidad que este se presente durante una operación por acciones no ilícitas. Este caso puede ser estudiado posteriormente y está por fuera del alcance de esta tesis.

4.4.2 Evaluación de aislamiento de L/G en segmentos m_{ij} por DS

El método de fusión por Dempster-Shafer realiza la combinación de información de mayores instancias dado que utiliza la información de un intervalo de confianza mayor a ± 3 desviaciones estándar. Esta posibilidad difiere ampliamente de la fusión por Bayes dado que los intervalos de confianza son mayores y pueden alcanzar un 95% al utilizar la información contrastada de tres desviaciones estándar.

Manteniendo el mismo esquema de pruebas que se hicieron para el método de fusión de Bayes, se tomará una pérdida/ganancia de un 1% para el modelo de segmento m_{ij} plasmado en la figura 4.19. Los resultados se plasmarán en la figura 4.28.

Figura 4-28: Fusión de datos con referencia terminales con L/G 1% por DS¹⁹.



Según los resultados dados por la fusión de la información por Dempster-Shafer del segmento m_{ij} se evidencia lo siguiente:

- Aunque las variaciones de la fusión por Dempster-Shafer son menores que las hechas por Bayes, las diferencias contrastadas con las pérdidas/ganancias son apreciables relativamente a la escala. Una estrategia de umbral estático es aplicable para detectar L/G en segmentos m_{ij} .
- El límite de fusión para los datos sin pérdida es 1. Esta probabilidad es lógica a comparación de los resultados dados para la fusión por Bayes. Esto garantiza una reciprocidad con el modelo aplicado para su análisis posterior. Dada la agrupación de la información de los valores fusionados, las desviaciones son evidentes pero en una escala de valores menores del 20%.

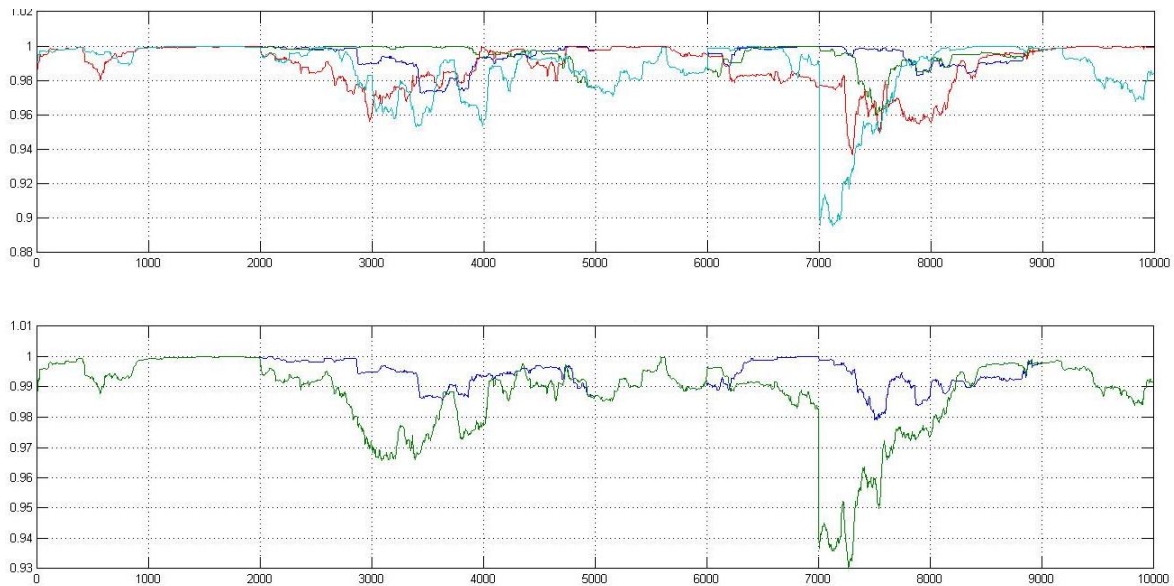
¹⁹ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Dempster-Shafer para una L/G del 1% con medidores de referencia terminales.

- La velocidad de reacción ante la pérdida o ganancia es menor a las 1000 muestras simuladas necesarias para evidenciar una L/G. Esto apoyaría ampliamente la oportunidad de la información y reacción de esta ante las desviaciones que se presentarán en la operación de un DTHR.
- El efecto de amortiguamiento se reduce al utilizar este método de fusión dado que la variación es menor que utilizando el método de fusión por Bayes. Este amortiguamiento se presentaba cuando se generaban las probabilidades tomadas con el muestreo de factor K y variaba dependiendo del tamaño del mismo factor.
- Los valores de fusión contrastada contra el medidor inicial tiene una mayor magnitud que las contrastadas por el medidor final. Esta diferencia radica en similitud estadística entre el medidor m_j y el medidor m_{fin} en comparación al medidor m_i .

Se utilizará la estrategia tomando como referencia medidores interlínea para ser fusionados por Dempster-Shafer utilizando la misma estrategia que se tomó para la fusión por Bayes. Los resultados se presentaran en la figura 4.29.

Analizando los resultados presentados en la figura 4.29, se pueden concluir diferentes aspectos con diferencia al análisis tomado en cuando los medidores de referencia son los terminales. Aún se evidencia variaciones por las L/G simuladas a las observaciones aunque el impacto es menor en probabilidad que los vistos para medidores terminales. Tanto como las velocidades de reacción y la longitud de estas variaciones no varían con respecto a los medidores terminales. Esta fusión es aplicable para el aislamiento de segmentos con pérdidas/ganancias para puntos de referencia de medidores terminales e interlínea.

Figura 4-29: Fusión de datos con referencia interlínea con L/G 1% por DS²⁰.



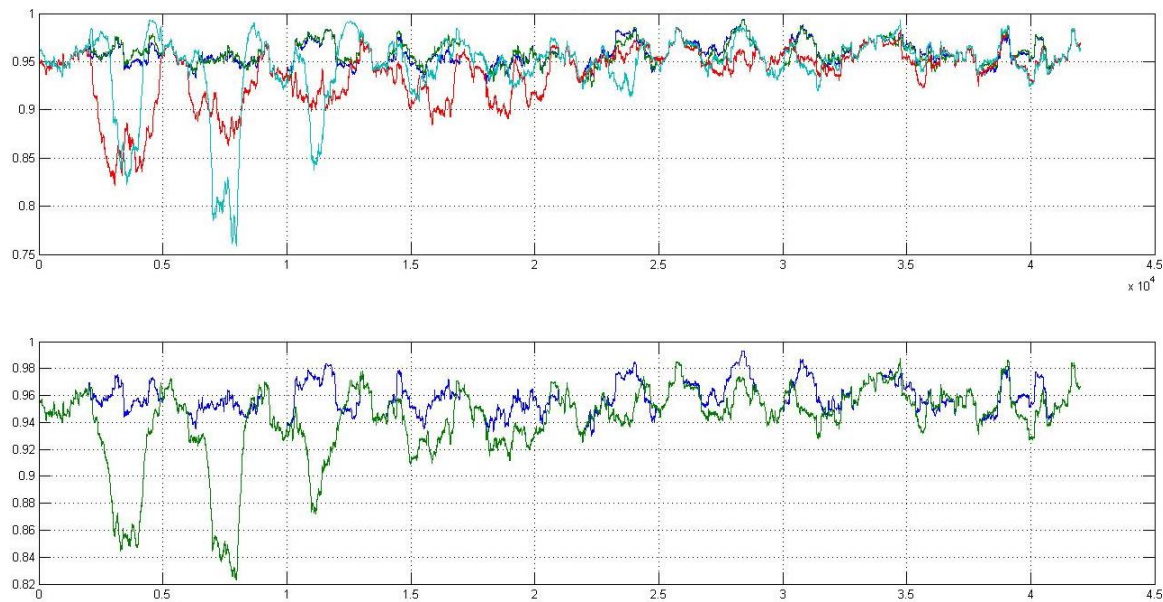
Sensibilidad de L/G para fusión DS para el segmento m_{ij} .

Con estas pruebas se definirá de una manera simple el nivel de sensibilidad ante las variaciones de desviaciones de L/G en un segmento m_{ij} de un DTHR mediante el uso de la fusión por Dempster-Shafer. Las amplitudes de las desviaciones valoradas por este método de fusión aunque son de menor magnitud que las hechas por el método Bayes son mejor diferenciables ante la detección de ratas de flujo bajas de desviaciones de pérdida/ganancia. Esta propiedad característica deberá permitir una mejor interpretación de las respuestas ante la detección de desviaciones L/G de baja rata.

Se utilizará el marco de pruebas hecho para probar la sensibilidad de fusión para el método Bayes para mantener la homogeneidad y hacer un análisis comparativo entre ambos métodos posteriormente. Dada la criticidad y eventualidad de tener en un momento determinado múltiples L/G se verificará también el uso de medidores de referencia terminales e interlínea. Teniendo como referencia los medidores interlínea, los resultados se presentan en la figura 4.30.

²⁰ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Dempster-Shafer para una L/G del 1% con medidores de referencia interlínea.

Figura 4-30: Fusión para un segmento m_{ij} por DS con L/G del 1% al 0,1% T²¹.

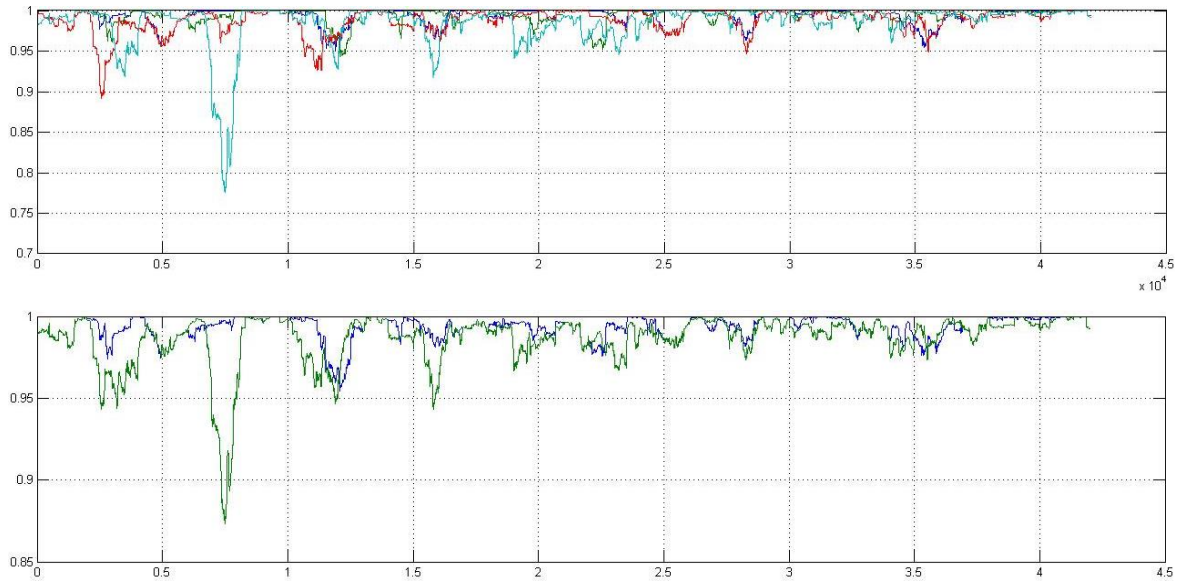


De los resultados de sensibilidad ante las ratas de flujo pequeñas de L/G usando el método de fusión de Dempster-Shafer se concluye lo siguiente:

- Las variaciones de los valores de fusión se mueven entre 100% y 92% cuando no existen pérdidas. Esto es repetitivo en las simulaciones hechas para la comprensión de la fusión por este método.
- La sensibilidad del método es observable hasta llegar a una pérdida/ganancia del 0.5%. Este valor de sensibilidad es igual al determinado para el método Bayes y se define que se puede utilizar ambos métodos indistintamente para el aislamiento de segmentos m_{ij} . Esta aseveración aplicará para el uso de umbrales de estado ante las variaciones, otros métodos podrían ser utilizados para la detección de estas desviaciones.
- Se evidencia una respuesta clara en las pérdidas de entre 1% al 0.97% de pérdidas pero posterior a estos valores el decaimiento es rápido y poco diferenciable para detectar alguna desviación.

²¹ Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Dempster-Shafer para varias L/G del 1% al 0,1% con medidores de referencia terminales.

Figura 4-31: Fusión para un segmento m_{ij} por DS con L/G del 1% al 0,1% ²².



En la gráfica 4.31 se puede observar el comportamiento de la sensibilidad de la detección por fusión de Dempster-Shafer tomando como referencia el medidor interlínea. Se puede observar un comportamiento similar a la respuesta que se tuvo con el método de Bayes, la detección es observable hasta una pérdida del 0.6% después de esto el comportamiento se mezcla con las oscilaciones operativas de la incertidumbre de los medidores.

Aislamiento de única L/G por fusión Dempster-Shafer para múltiples segmento m_{ij}

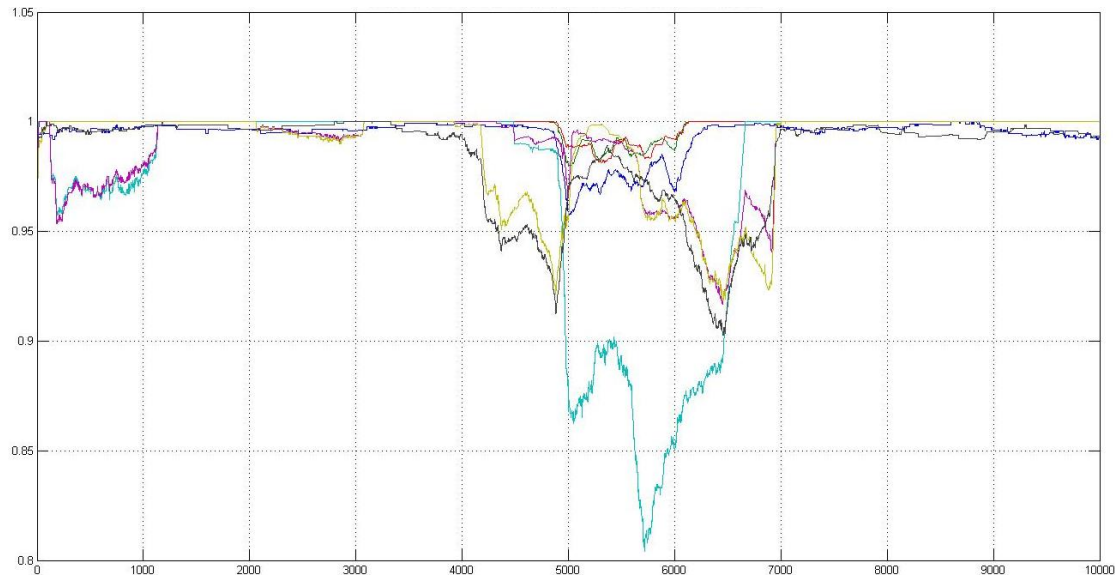
Partiendo de las pruebas hechas para probar la fusión por Bayes en aislamiento de segmentos se probarán para Dempster-Shafer. Inicialmente se probará una pérdida en el segmento m_{45} por un periodo de 2000 momentos y con medidores de referencia terminales. Se modificará la matriz de covarianza del filtro Kalman de las observaciones Kalman a 0.00001 dada la sensibilidad de método a utilizar y su respuesta. Los resultados están presentados en la figura 4.32.

Para maximizar la interpretación del análisis hecho por el método Dempster-Shafer para el aislamiento y las características probabilísticas de los valores entregados posteriormente por este método, se multiplicarán los valores de las fusiones hechas

²² Resultados de fusión de datos de medidores de flujo para un segmento m_{ij} por fusión de Dempster-Shafer para varias L/G del 1% al 0,1% con medidores de referencia interlínea.

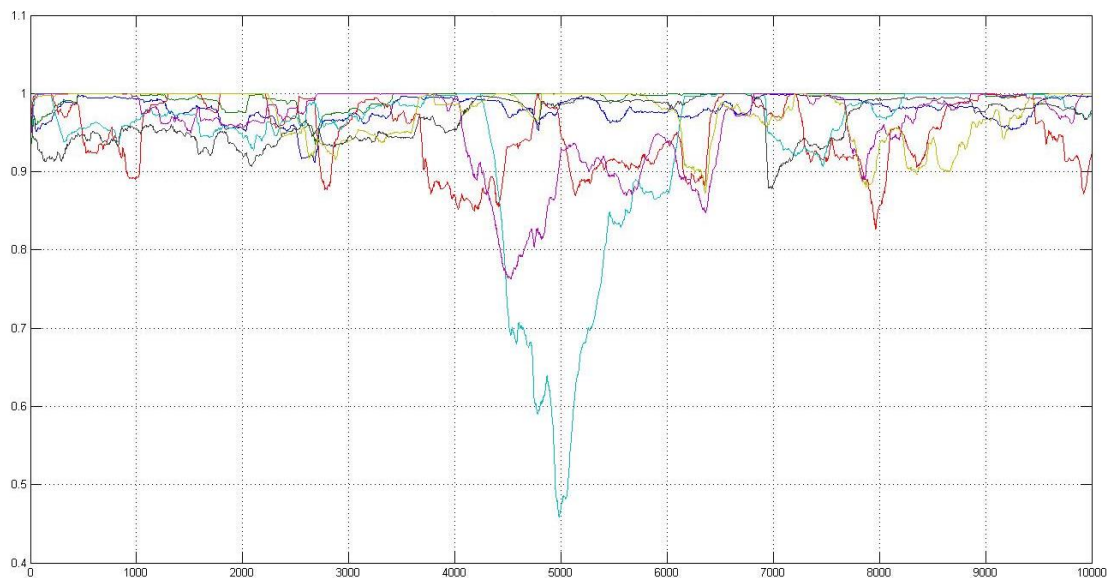
contra los medidores terminales. Esto permitirá observar e interpretar las tendencias de dicho método.

Figura 4-32: Fusión datos múltiples segmentos m_{ij} por DS, única L/G de 1% T.



La respuesta dada por el método es notoria a la respuesta de las fusiones hechas para los segmentos que no presentan esta desviación. Se puede notar hasta de un 0.7 de la probabilidad para el segmento m_{45} para una pérdida de un 1%. Se puede observar que los demás segmentos presentan una afectación por la pérdida, esto permitirá posteriormente diferenciar y clasificar una desviación de L/G de otros tipos de fallos, como pérdida de calibración de instrumentos de medición o fallas internas de los mismos.

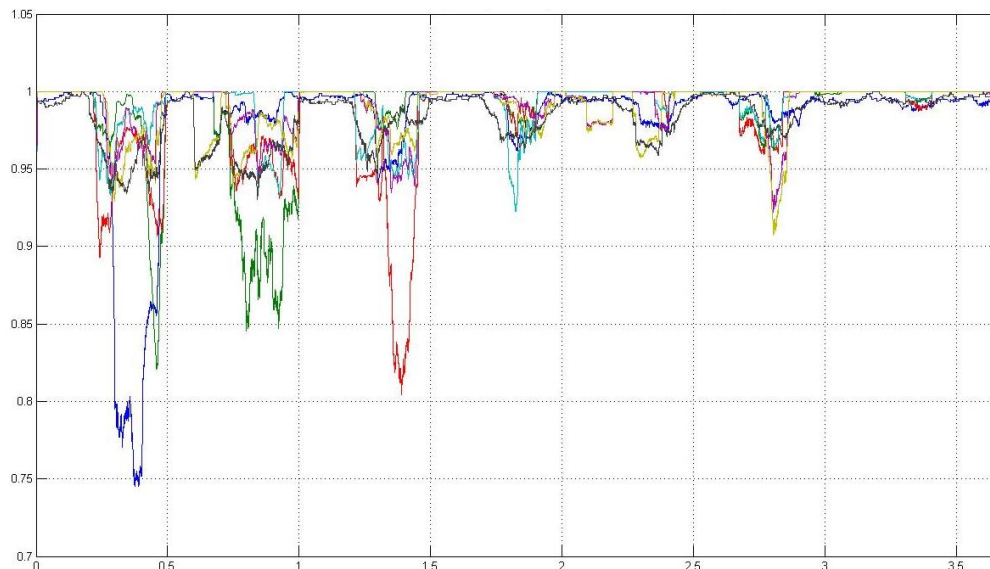
Figura 4-33: Fusión datos múltiples segmentos m_{ij} por DS, única L/G de 1% I.



Para el caso en que se utilice como referencia los medidores interlínea adyacentes al segmento m_{ij} los resultados se presentan en la figura 4.33. La variaciones en las observaciones hechas tienen mayor amplitud que cuando se utilizó contrastando con medidores terminales debido a la naturaleza aleatoria de las mismas. En este caso el comportamiento de las fusiones de los otros segmentos no presenta un comportamiento inestable como en la anterior prueba.

Utilizando el mismo marco de prueba para la fusión con Bayes que introdujo diferentes pérdidas en segmentos con ratas diferentes se utilizará para la fusión con Dempster-Shafer. Esta prueba se plasmará en la figura 4.34

Figura 4-34: Fusión datos múltiples segmentos m_{ij} por DS, múltiples L/G de 1% T²³.



Se puede observar que el método de fusión de Dempster-Shafer tiene una respuesta similar a la fusión de Bayes en función de la sensibilidad, tiene un aislamiento cercano a 0,5% de desviación utilizando la ecuación 4.3. Se pueden detectar desviaciones hasta 0.5% lo cual es limitado a comparación con el método Bayes.

$$\min \left(\Delta M_{ij} = \sum_{r_0}^{r_n-K} \widehat{m}_{ij_{ini}} \times \widehat{m}_{ij_{fin}} \right) \quad (4.3)$$

4.5 Caracterización y cuantificación de desviaciones L/G (nivel 4)

Esta sección pretende demostrar que con el uso de los medidores de flujo independientemente de su incertidumbre pueden observar y cuantificar con cierta precisión las pérdidas o ganancias en sectores donde se han identificado y aislado pérdidas previamente. Esta sección parte de las mediciones observadas y fusionadas por

²³ Resultados de fusión de datos de medidores de flujo para múltiples segmentos m_{ij} por fusión Dempster-Shafer para múltiples L/G en diferentes sectores con medidores de referencia terminales.

método de ponderados de confianza, filtradas o no, y de los resultados hechos por los niveles 2 y 3, identificación y aislamiento de L/G. EL resultado de esta sección presentará el método más eficaz para cuantificar las pérdidas y ganancia.

Para generar una comparación apropiada con la metodología hecha para los niveles 2 y 3, se tomará un DTHR con 6 medidores interlínea y dos terminales. Este modelo permitirá mostrar los resultados que se tendrán de la medición de las L/G inducidas por simulación.

Tomando solo 20 muestras de una pérdida de 1% sobre las pruebas propuestas en la metodología en el segmento 4&5 se presentarán los resultados en la tabla 4.9.

Tabla 4-9: Caracterización de L/G de pérdida de 1% en el segmento 4&5.

Muestra	Sust. Sin F	Desv.	Sust. Con F	Desv.2	Fusión M	desv.3
1	1,0981	-0,098%	1,0936	-0,094%	1,0871	-0,087%
2	0,9280	0,072%	0,9256	0,074%	0,9241	0,076%
3	0,9502	0,050%	0,9470	0,053%	0,9642	0,036%
4	0,9786	0,021%	0,9778	0,022%	0,9593	0,041%
5	0,9171	0,083%	0,9141	0,086%	0,9517	0,048%
6	0,9349	0,065%	0,9334	0,067%	0,9415	0,059%
7	0,9346	0,065%	0,9305	0,070%	0,9303	0,070%
8	0,9972	0,003%	0,9918	0,008%	0,9772	0,023%
9	1,0512	-0,051%	1,0516	-0,052%	1,0479	-0,048%
10	1,0235	-0,024%	1,0252	-0,025%	1,0027	-0,003%
11	1,1301	-0,130%	1,1300	-0,130%	1,1328	-0,133%
12	1,0048	-0,005%	1,0071	-0,007%	0,9992	0,001%
13	1,0984	-0,098%	1,1005	-0,101%	1,1046	-0,105%
14	0,9206	0,079%	0,9237	0,076%	0,9524	0,048%
15	1,0200	-0,020%	1,0127	-0,013%	1,0371	-0,037%
16	0,9021	0,098%	0,8985	0,102%	0,9021	0,098%
17	0,9841	0,016%	0,9766	0,023%	0,9655	0,035%
18	0,9870	0,013%	0,9868	0,013%	0,9886	0,011%
19	0,9830	0,017%	0,9854	0,015%	0,9796	0,020%
20	0,8620	0,138%	0,8642	0,136%	0,8727	0,127%
Promedio	0,9853		0,9838		0,9860	
Máximo	1,1301		1,1300		1,1328	
Mínimo	0,8620		0,8642		0,8727	
Desv.Est	0,0706		0,0707		0,0671	
Varianza	0,0050		0,0050		0,0045	
Curtosis	2,5255		2,5115		2,8252	

Ahora tomando 20 muestras de una ganancia de 1% sobre las pruebas propuestas en la metodología en el segmento 4&5 se presentarán los resultados en la tabla 4.10.

Tabla 4-10: Caracterización de L/G de ganancia de 1% en el segmento 4&5.

Muestra	Sust. Sin F	Desv.	Sust. Con F	Desv.2	Fusión M	desv.3
1	-1,0250	-0,025%	-1,0302	-0,030%	-1,0182	-0,018%
2	-1,0401	-0,040%	-1,0361	-0,036%	-1,0037	-0,004%
3	-0,8777	0,122%	-0,8826	0,117%	-0,8892	0,111%
4	-0,9120	0,088%	-0,9065	0,094%	-0,9123	0,088%
5	-0,7862	0,214%	-0,7968	0,203%	-0,8153	0,185%
6	-1,0436	-0,044%	-1,0476	-0,048%	-1,0634	-0,063%
7	-0,9418	0,058%	-0,9464	0,054%	-0,9488	0,051%
8	-1,0187	-0,019%	-1,0148	-0,015%	-1,0376	-0,038%
9	-0,9763	0,024%	-0,9673	0,033%	-0,9625	0,038%
10	-1,0301	-0,030%	-1,0228	-0,023%	-1,0083	-0,008%
11	-1,0174	-0,017%	-1,0144	-0,014%	-1,0130	-0,013%
12	-0,9204	0,080%	-0,9185	0,082%	-0,9163	0,084%
13	-1,0032	-0,003%	-0,9995	0,000%	-1,0182	-0,018%
14	-0,9951	0,005%	-0,9933	0,007%	-1,0209	-0,021%
15	-0,9655	0,035%	-0,9613	0,039%	-0,9402	0,060%
16	-1,1660	-0,166%	-1,1646	-0,165%	-1,1290	-0,129%
17	-1,0024	-0,002%	-1,0003	0,000%	-0,9812	0,019%
18	-0,7743	0,226%	-0,7685	0,232%	-0,7976	0,202%
19	-1,0500	-0,050%	-1,0491	-0,049%	-1,0661	-0,066%
20	-0,9276	0,072%	-0,9300	0,070%	-0,9484	0,052%
Promedio	-0,9737		-0,9725		-0,9745	
Máximo	-0,7743		-0,7685		-0,7976	
Mínimo	-1,1660		-1,1646		-1,1290	
Desv.Est	0,0912		0,0901		0,0819	
Varianza	0,0083		0,0081		0,0067	
Curtosis	3,5267		3,5726		3,0022	

Partiendo de la notación para definir pérdidas con signo positivo y ganancias con el signo negativo que se direcciona por API, se puede analizar de las tablas 4.9 y 4.10 lo siguiente:

- Cualquiera de los tres métodos utilizados es útil para cuantificar una pérdida/ganancia de 1%. Las diferencias no son sustanciales entre ellos y dado esto lo más práctico sería utilizar el método de resta directa dado que utiliza mucho menos recurso informático para establecer y caracterizar la L/G.

- Se nota una desviación estándar cerca al 8% entre las pruebas hechas, es decir, se puede llegar a una incertidumbre del mismo valor cuando se utiliza el método de fusión de medidores. La fusión para estas muestras, y otras hechas no plasmadas, presentaron una unidad menor que los otros métodos. Esta solución es útil para el caso en que la pérdida debe ser calculada con una máxima aproximación aunque no se garantice una precisión de custodia como tal.
- Si se cuenta con la información totalizada de los mismos medidores registrada con tiempo mejoraría ostensiblemente esta medición. Esta solución mezclada con la identificación y aislamiento de desviaciones L/G reduciría la carga informática del sistema y aumentaría la confiabilidad de la metodología aunque aumentaría la carga de información comunicada por cada medidor.
- Las otras variables estadísticas presentan un comportamiento parejo entre ganancias y pérdidas. Esto evidencia un comportamiento parejo para ambas caracterizaciones y no se agregaron no linealidades al cambiar el signo en el muestreo.

5. Otras condiciones operativas

En la operación de ductos de hidrocarburos se pueden tener escenarios comunes que afectarían la operatividad de la estrategia que se está planteando en la metodología propuesta. Ya se evaluó la metodología donde los escenarios operativos eran normales; toda la información de las mediciones estaba completa, coherente y sincronizada en tiempo. Este escenario operativo es común en los ductos dado la confiabilidad de los equipos de medición y la disponibilidad de las comunicaciones a los diferentes puntos donde se encuentran las mediciones. En este capítulo se tratará la aplicabilidad de la metodología propuesta en escenarios anormales de la operación para validar desde un punto de vista coherente la aplicabilidad de la misma en condiciones adversas.

Las condiciones anormales de operación que se tratarán en este capítulo se enmarcan en las siguientes: Desconexión, descalibración, congelamiento y desincronización de señales de equipos de medición. Estos escenarios operativos anormales se simularán para fallas individuales de los equipos de medición debido a que por los esquemas de confiabilidad que prestan los medidores y comunicaciones sería complejo tener ambas consideraciones al mismo tiempo. Fallas simultáneas podrán estudiarse posteriormente al desarrollo de esta metodología.

Cada condición operativa se validará en función de la identificación, aislamiento y caracterización de las pérdidas o ganancias que se tendrían. Esta validación permitirá comparar de una manera metodológica cada escenario propuesto para ser evaluado posteriormente bajo consideraciones de robustez de la metodología contra escenarios adversos de la operación real. Todos los casos evaluados se tendrán las mismas consideraciones de flujo y falla, es decir, se simulará a una rata de flujo de 100 unidades por periodo y la falla situada en el segmento 4 y 5 de la gráfica 4.20 reducida a dos medidores terminales y seis interlínea.

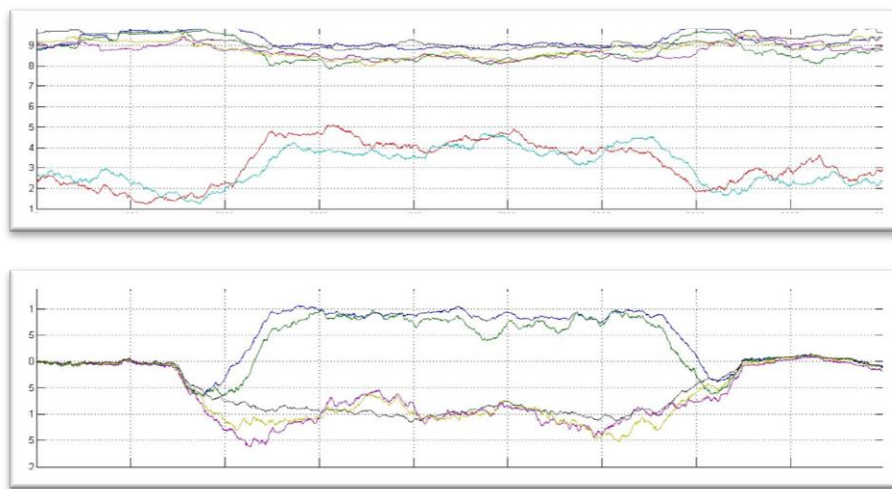
Se realizaran fusiones por Bayes y Dempster-Shafer de la información simulada para cuando se tiene como punto de referencia los medidores terminales e interlínea. Ambos análisis son importantes dado que una falla de un medidor de un segmento cualquiera permitiría ser tomado por otro medidor adyacente que solo por la información de los equipos terminales no se tendría y para que esto suceda es necesario tener ambos enfoques de análisis.

5.1 Desconexión de medidor de flujo

Una falla de desconexión de un medidor de flujo se presenta ya sea por la desconexión física del medidor a la fuente o por falta del suministro de energía, que es común en los sectores lejanos donde se encuentran estos. Una falla por suministro en la mayoría de los casos es de corta duración cuando se presenta ya que en los puntos de medición se cuenta con suministro alterno de alimentación de corto plazo como UPS y baterías. En el caso de una falla por desconexión física del elemento puede llevar más tiempo dado el desplazamiento de personal técnico a dicha reparación.

Una falla por desconexión impacta directamente la estrategia propuesta por falta de punto de referencia de comparación. La medición puntual afectada quedaría desvinculada de la realidad operativa del ducto provocando que el análisis comparativo del segmento pierda validez frente al manejo de los otros sectores.

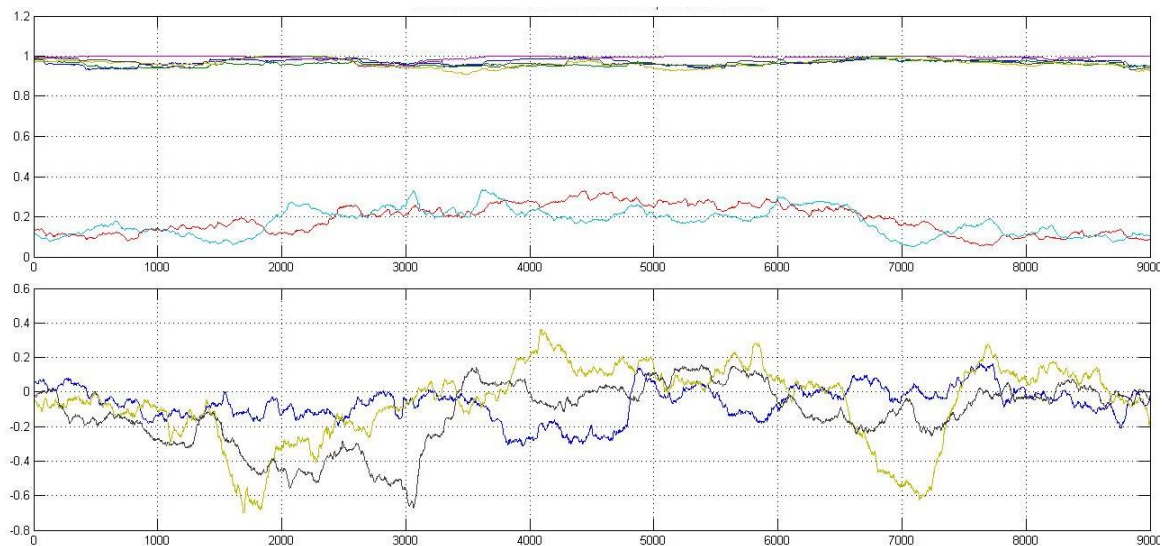
Figura 5-1: Falla de desconexión con referencia medidores terminales



Evaluando la falla de desconexión cuando se presenta una pérdida de 1 unidad por periodo se presenta en la figura 5.1 con referencia en los medidores terminales. Evaluando según criterios de la metodología propuesta se tendría lo siguiente:

- *Identificación:* Se identifica una pérdida en el sistema con claridad por ambos métodos de fusión. Para el método Dempster-Shafer se observa que los resultados de la fusión del segmento en falla trata de acercarse cuantitativamente a los de los resultados de los segmentos en operación normal, esta consideración es lógica debido a que la falla representaría un número 0 que está más cercano al 99 que el 100 en términos de probabilidad. Para la fusión Bayes los datos son nulos dado que existe una multiplicación aritmética por cero dentro de la fusión.
- *Aislamiento:* No es posible determinar qué segmento presentó la pérdida analizándolo por ambos métodos de fusión. Mientras que la información proveniente de la fusión de Bayes no existe, para la fusión hecha por Dempster-Shafer no es determinante para asignar la pérdida a dicho sector.
- *Caracterización:* A la falta de información de este medidor, no es posible realizar una cuantificación o naturaleza de la desviación operativa.

Figura 5-2: Falla de desconexión con referencia medidores interlínea.



En la figura 5.2 se plasman los resultados de la falla a desconexión cuando existe referencia a medidores interlínea. La respuesta es lógica dado que la dependencia de los medidores en este análisis es mucho mayor que en el análisis con referencia de medidores terminales. Evaluando según los criterios:

- *Identificación:* No existe una respuesta clara de haya una desviación operativa bajo el modelo de análisis con referencia interlínea. Ambas fusiones requieren obligativamente la información de los medidores interlínea y al desaparecer la información fusionada es nula.
- *Aislamiento:* Se evidencia por ambos métodos la falla pero no la desviación operativa por segmento. Esta consideración puede diagnosticar el estado de los medidores de flujo pero no las pérdidas o ganancias al evidenciar que las respuestas de fusiones adyacentes al medidor en falla son nulas.
- *Caracterización:* Idéntica respuesta al análisis hecho para la misma falla con medidores de referencia Terminales.

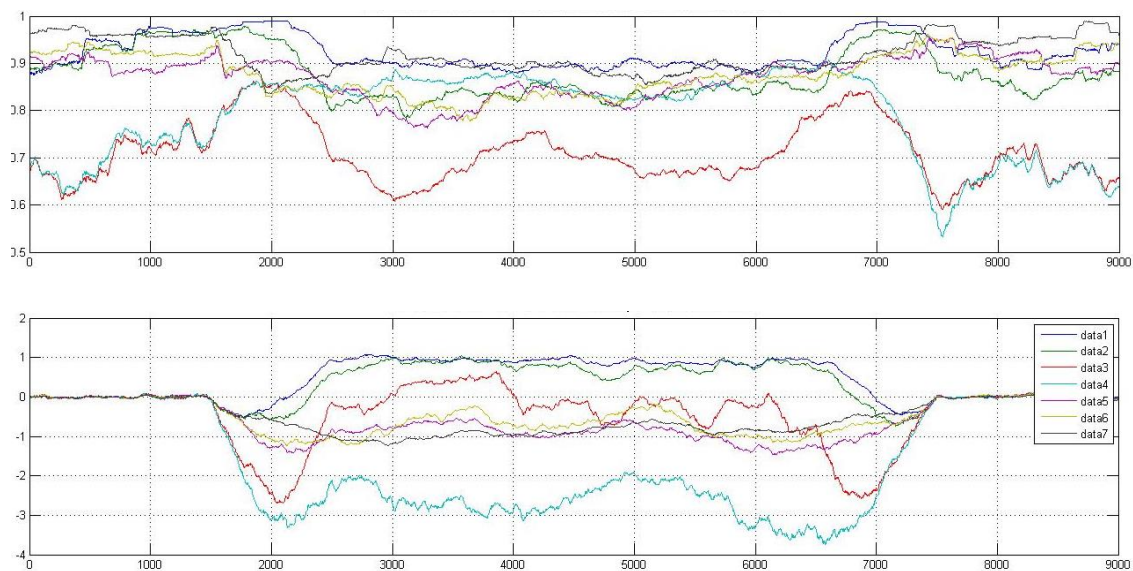
Una falla en desconexión en cualquier medidor terminal en ambos análisis ocasionaría la inoperancia del método propuesto en esta tesis. Esta falla lógicamente afectaría en una mayor medida al análisis hecho basado en la información suministrada por el mismo medidor. Esto aunque es viable que suceda, el esquema de operación de los ductos no les permite operar sin este medidor dado las consideraciones de control y riesgo de hacerlo sin él.

5.2 Descalibración del medidor de flujo

La descalibración de equipos de medición de flujo es frecuente en la industria e inherentes a su naturaleza y operación. Para contrarrestar esta situación, las industrias del petróleo cuentan con equipos y personal de metrología para garantizar el funcionamiento preciso y estadísticamente adecuado de los diferentes medidores. Dado que las frecuencias de reparación pueden ser pocas al año dependiendo de los esquemas de mantenimiento, es necesario evaluar la incidencia de esta situación en la metodología que se está proponiendo.

Para evaluar esta situación, se llevará al medidor m_4 con una desviación de un 1% de la medida real de flujo y se simulará una desviación operativa a pérdida de un barril por hora. Se prevé un impacto menor que la falla a desconexión debido principalmente a la descripción de la situación operacional esta levemente desfasada de la realidad. Se hará el ejercicio para las fusiones de BY y DS con mediciones de referencia terminal e interlínea.

Figura 5-3: Falla de descalibración con referencia medidores terminales.

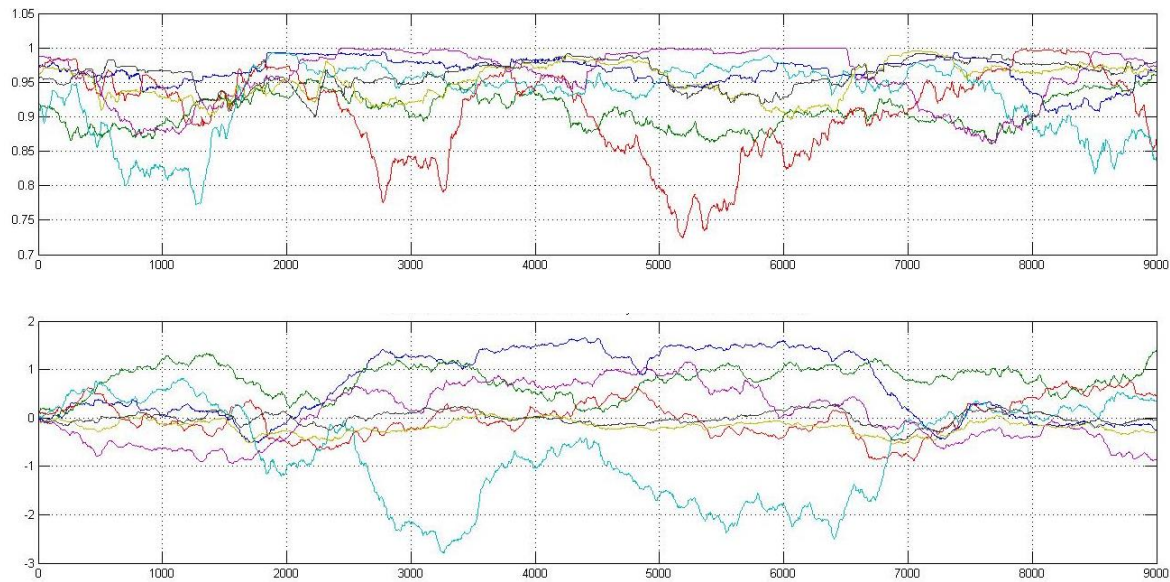


Teniendo como referencia que la respuesta de la descalibración del medidor cuatro lo referencian las tendencias rojas (TR) y aguamarina (TA), se evalúa con los criterios predefinidos:

- *Identificación:* Se evidencia una desviación operativa por ambos métodos de fusión. Por DS se puede evidenciar las fallas de descalibración en las TA y TR inmediatamente anteriores y posteriores a la desviación operativa analizada. Dado que estas dos fusiones tiene como punto común el medidor cuatro, es sencillo establecer la falla del mismo bajo este método. Para la fusión BY no se presenta en los momentos anteriores o posteriores esta condición.

- *Aislamiento:* A pesar que la evidencia de una desviación operativa es clara y se trató en la identificación, el aislamiento del segmento en falla no es sencillo su separación del comportamiento de los otros.
- *Caracterización:* Las mediciones hechas bajo las estrategias de caracterización sugeridas en el capítulo cuatro van de valores desviados hasta un 168% para la desviación operativa inducida.

Figura 5-4: Falla de descalibración con referencia medidores interlínea.



Siguiendo las mismas notaciones para las TR y TA, se puede hacer los siguientes análisis para los criterios a evaluar:

- *Identificación:* Para las dos tendencias a analizar en la fusión DS existe una desviación operativa para la TR de la fusión BY no se puede decir lo mismo. Esto parte del hecho que al tener la misma magnitud la descalibración con la desviación operativa estadísticamente se “escondería” dicha información dentro del periodo de la misma.
- *Aislamiento:* la fusión DS presenta una respuesta clara de aislamiento aunque en los momentos previos y posteriores de la desviación operativa se puede notar que la TA presenta un comportamiento aislado del resto de tendencias. Para la fusión BY e igual al análisis hecho para la identificación, no se podría dar una conclusión total del aislamiento del segmento.

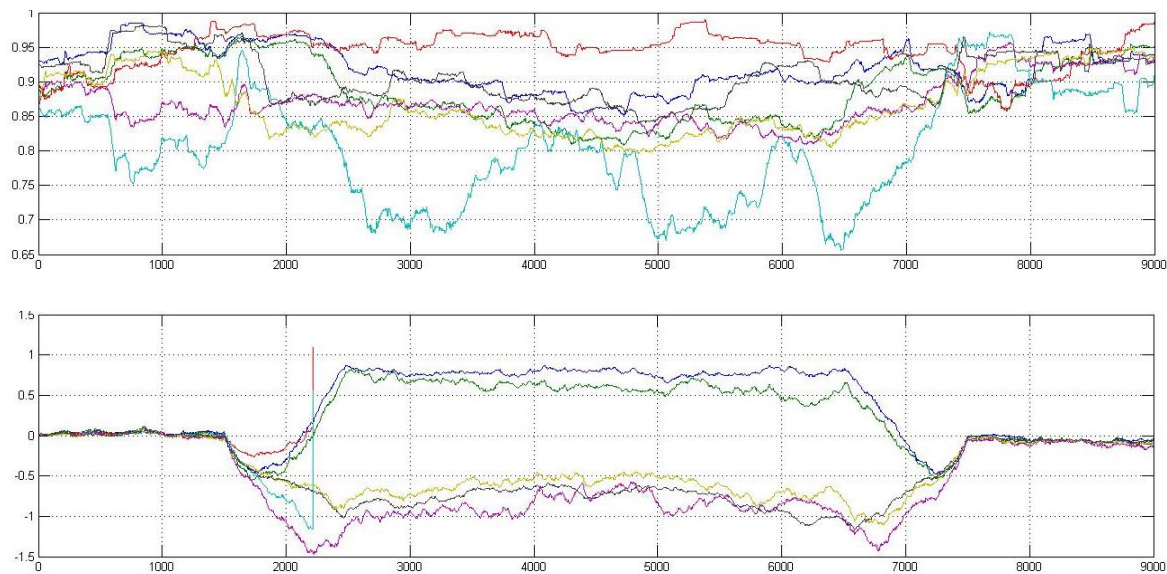
- *Caracterización:* Se tiene una desviación promedio para mil iteraciones de 221% de desviación a la pérdida según el método de resta aritmética con filtro Kalman. Hay un claro impacto de esta falla en la caracterización de las pérdidas.

5.3 Congelamiento de la señal del medidor de flujo

Una falla por congelamiento se presenta cuando la señal enviada por el medidor de flujo no presenta cambio en un periodo de tiempo determinado. Las causas varían siendo la principal fallas de transmisión en comunicación por la actualización del nodo o una falla interna del medidor que entrega esta información. Esta última falla es poco frecuente debido a los esquemas de autodiagnóstico de los equipos actualmente pero la falla por comunicación es frecuente en sistemas SCADA. Para este caso en particular, se asignara un valor estático de 100 unidades por periodo al medidor cuatro.

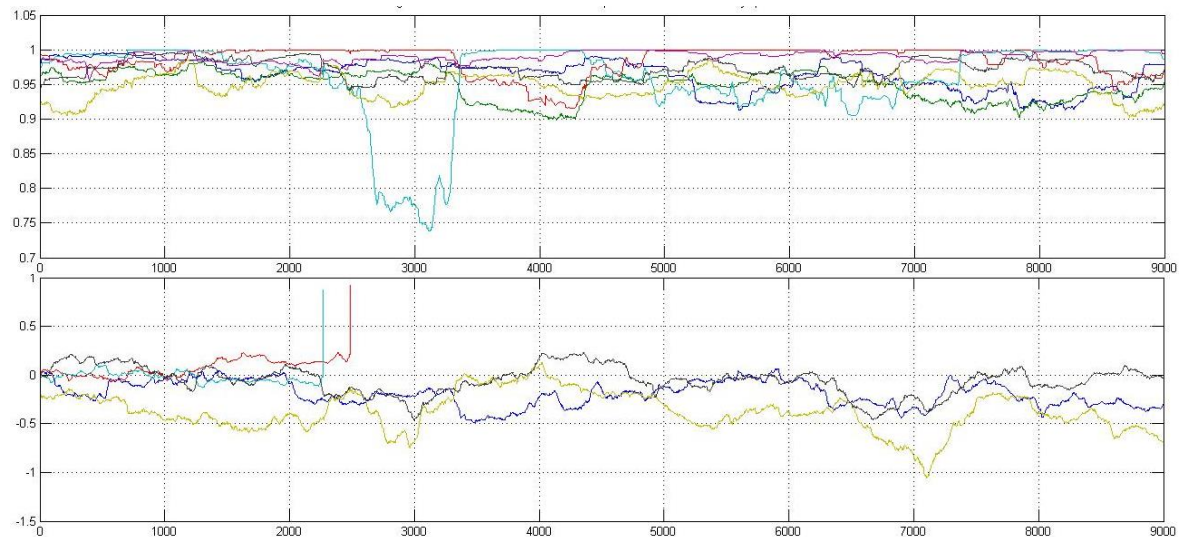
Esta falla afecta las fusiones hechas de información dado que la condición de no aleatoriedad de información sumado a la no coherencia con la realidad operativa generar que las respuestas de los cálculos matemáticos se saturen. Esta condición podría generar una estrategia adicional para chequeo de la información en condiciones normales y pérdida. Se generaría un medidor virtual con un dato estático del valor promedio del medidor inicial en un periodo de desviación más un valor de ganancia o pérdida estimada, esto aumentaría la respuesta de las fusiones. Esta estrategia se podrá evaluar en investigaciones posteriores.

Figura 5-5: Falla por congelamiento con referencia medidores terminales.



Siguiendo las mismas notaciones para las TR y TA, se puede hacer los siguientes análisis para los criterios a evaluar:

- *Identificación:* Ambas estrategias de fusión responden a la pérdida inducida al sistema. Para la fusión BY se puede notar que las TR y TA se saturan momentos posteriores a la desviación inducida; esto es dado a que las correcciones aritméticas anteriores son muy cercanas al cero y al aproximar estas quedan nulas impactando las correcciones posteriores. La fusión DS no presentó la falla de la fusión BY para este caso.
- *Aislamiento:* Con la fusión BY no se puede dar un veredicto final en donde se presentó la desviación operativa debido a la saturación de la respuesta de la misma. Para la fusión DS, la TR no presenta una respuesta representativa a la desviación operativa, en cambio la TA si manifiesta una desviación. Esto es debido básicamente a la que la TA depende más del medidor cinco que si evidencia una desviación importante contra lo presentado en el medidor cuatro.
- *Caracterización:* Se encontraron desviaciones menores del 2% con la fuga incidente en el sector. Dado que el medidor cuatro se mantuvo fijo al flujo promedio se pudo establecer claramente con la señal filtrada del medidor cinco.

Figura 5-6: Falla por congelamiento con referencia medidores interlínea.

Siguiendo las mismas notaciones para las TR y TA, se puede hacer los siguientes análisis para los criterios a evaluar:

- *Identificación:* Por fusiones de BY y DS ya no es posible la identificación de la desviación operativa inducida a esta. Se evidencia una respuesta solitaria por DS en la TA pero no es concluyente para un análisis concienzudo.
- *Aislamiento:* Se evidencia una saturación de la fusión BY para ambos sectores involucrados como paso en el análisis hecho para el uso de referencia terminal. Para la fusión DS se evidencia una saturación continua a uno de la TR, TA manifiesta una desviación pero no es concluyente.
- *Caracterización:* No está claro el periodo de las desviaciones y por tanto no se puede definir la cuantía y la naturaleza de la desviación.

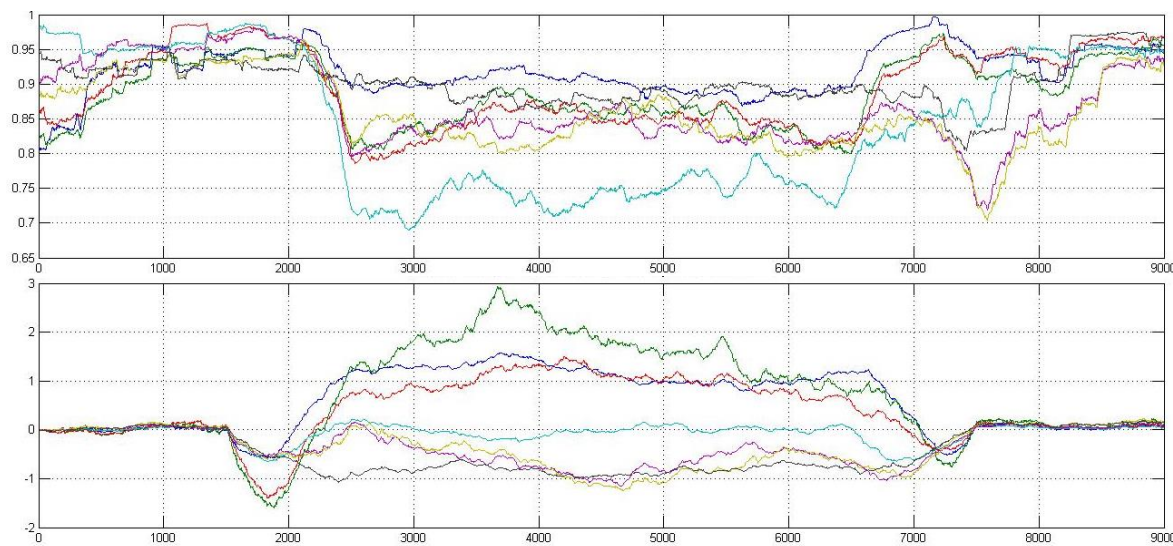
5.4 Desincronización de información de los medidores

Esta falla es frecuente para sistemas distribuidos cuando la información no llega en el tiempo o no puede ser organizada de manera temporal. Esta falla puede ser mitigada con el uso de relojes con sincronización GPS en cada nodo o protocolos de comunicación hechos bajo estas premisas como NTP. Dada las consideraciones de distancia y ubicación de los diferentes nodos o puntos de medición de los ductos, esta falla puede

ser recurrente en la operación de los mismos si no se cuenta con sistemas de sincronización de tiempo y uso de estampas de tiempo “time-stamp”.

Esta falla es crítica para las fusiones de información dado que la metodología se basa en el análisis de los eventos en momentos definidos e idénticos para todos los medidores de flujo. Para este análisis se incidirá una desviación operativa a pérdida de 1 barril por hora en un momento k definido por un periodo K , el medidor cinco presentará la desviación en $k + n$ momentos y el medidor seis lo harán en $k + 2n$ momentos.

Figura 5-7: Falla por desincronización m_4 y m_5 con referencia terminal.

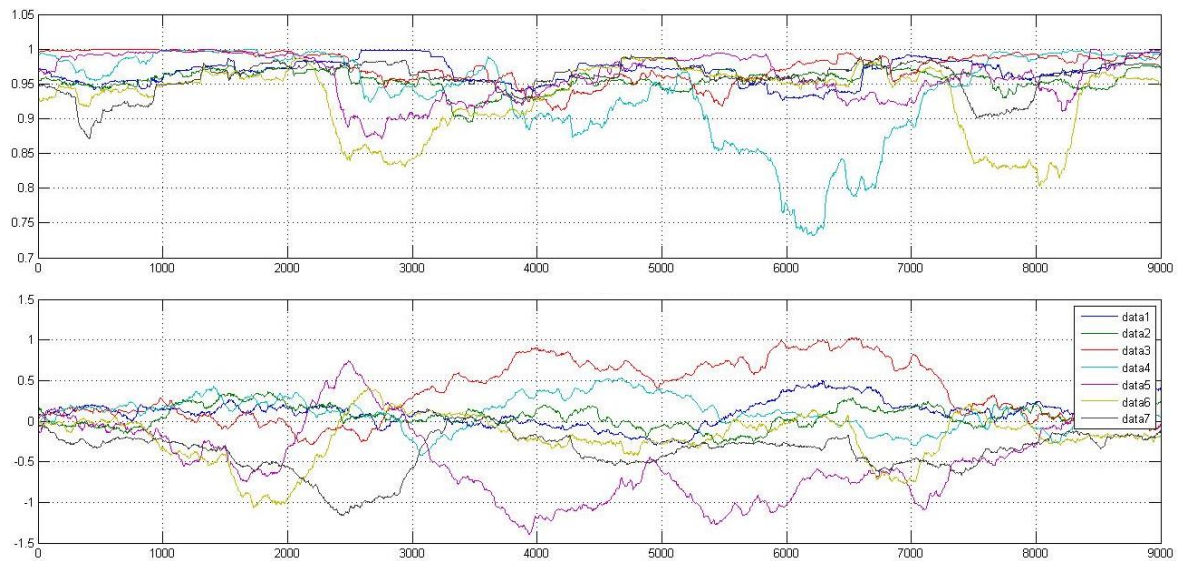


Siguiendo las mismas notaciones para las TR y TA, se puede hacer los siguientes análisis para los criterios a evaluar:

- **Identificación:** Ambas fusiones identifican la desviación operativa inducida. Para la fusión DS se evidencia una reducción en el impacto de la medición, es decir, para una condición normal este puede bajar hasta un 0,4 en la respuesta combinada, ahora este reportando un mínimo de 0,7 en TA.
- **Aislamiento:** Para fusión de BY es evidente el aislamiento de la desviación pese a desincronización de la información. Para la fusión DS es complejo dado que TR aunque identifica la desviación está inmersa en la respuesta de los otros segmentos, TA evidencia la desviación.

- **Caracterización:** Debido a que el periodo de la desviación identificada y aislada es más largo que la que se incidió inicialmente, la desviación de la medición se encontró en 183%.

Figura 5-8: Falla por desincronización m_4 y m_5 con referencia interlínea.



Siguiendo las mismas notaciones para las TR y TA, se puede hacer los siguientes análisis para los criterios a evaluar:

- **Identificación:** No se evidencia una desviación operativa siguiendo los métodos propuestos para esto por ambas fusiones.
- **Aislamiento:** No se puede determinar que segmento presenta la desviación operativa inducida por ambos métodos de fusión.
- **Caracterización:** Al no tener definido el periodo de tiempo de la desviación operativa y el segmento donde se presentó, no se puede cuantificar y definir naturaleza de la misma.

6. Conclusiones y recomendaciones

6.1 Conclusiones

- El uso de este método permitirá el aumento de la oportunidad en la respuesta de control para eventos de L/G hasta del 0,5% del flujo nominal en operación de DTHR, salvando vidas, protegiendo el medio ambiente y reduciendo efectos económicos potenciales a las empresas transportadoras.
- Al utilizar recursos de medición instalados a lo largo de los DTHR, la implementación de la metodología repercutiría en un menor costo de implementación que al instalar nuevo sistema de detección de fugas con igual rendimiento y mayor equipo asociado.
- El método propuesto permite una identificación, aislamiento y caracterización en conjunto de desviaciones operativas L/G con magnitudes hasta del 0.5% del flujo instantáneo sobre el DTHR. Esto se logra con sistemas de medición terminal e interlínea con incertidumbres del 0.5% y 5% del flujo instantáneo respectivamente.
- El filtro Kalman se adapta mejor que el filtro ponderado a las mediciones interlínea y condiciones operativas presentes en la operación de un DTHR tal como se modeló.
- Al aplicar el filtro Kalman a las señales entregadas por los medidores interlínea, la respuesta de las desviaciones operativas L/G al contrastar estos valores probabilísticamente con las mediciones hechas por los medidores terminales aumentará. Se evidenció que las señales filtradas por este método tienen un comportamiento homogéneo con las mediciones del medidor terminal.

- La identificación de las desviaciones operativas L/G dependen de la magnitud de la misma; usando el método propuesto puede evidenciar magnitudes menores a 0,1% del flujo instantáneo.
- Al aislar un segmento del DTHR por una desviación operativa L/G utilizar como referencia los medidores terminales en vez de los interlínea asegura una mejor definición de este.
- La efectividad del aislamiento depende directamente de la magnitud de la desviación operativa L/G.
- La fusión por Dempster-Shafer proporciona una mejor interpretación del aislamiento de segmentos de las desviaciones operativas L/G que la fusión Bayes tal como se modeló y con referencia ambas referencias; medidores terminales e interlínea.
- La caracterización de las desviaciones operativas L/G dependen directamente de la capacidad de determinar el periodo (identificación) y la segmentación (aislamiento) en que existió ésta. Los métodos de filtrado y fusión de la información no fueron más efectivos dada esta condición.
- La metodología de fusión de información propuesta para mejorar la incertidumbre de los medidores interlínea no fue aplicada en esta metodología por su bajo impacto en la topología del DTHR utilizada.
- Es necesario garantizar el adecuado funcionamiento de los medidores de flujo y sistemas de comunicación de un DTHR para que el método propuesto sea efectivo para la identificación, aislamiento y caracterización de desviaciones operativas L/G.
- El método propuesto presenta diferentes reacciones a las fallas modeladas permitiendo caracterizar cada una de ellas de una manera diferente.

- La fusión de Dempster-Shafer se evidencio una baja incidencia por algunas de las fallas modeladas, la fusión Bayes para el aislamiento de desviaciones operativas en condiciones de falla se vio afectada directamente.

6.2 Recomendaciones

- Aplicar técnica para manejo de umbrales de detección dependiendo de la operación de los DTHR. Se sugiere el uso de umbrales estáticos, dinámicos y por gradientes comparados.
- Se sugiere explorar la aplicación de este método en aplicaciones donde haya múltiples fuentes y único objetivo. Estas aplicaciones son frecuentes en esquemas de logística y transporte.
- Llevar este método en aplicaciones con topologías de distribución más complejas (mallas, estrellas, anillos, mixtas, entre otros) para la identificación, aislamiento y caracterización de desviaciones operativas L/G. Estas topologías son frecuentes en campo de producción petrolera, transporte y distribución de combustibles, entre otros.
- Verificar el uso de mediciones residuales entre medidores o virtuales para el manejo de referencias adicional a las referencias con medidores terminales e interlínea. Esta aplicabilidad es importante para la implementación en topologías más complejas.
- Utilizar la metodología para la identificación, aislamiento y caracterización de fallas en otras aplicaciones que conlleven sistemas con incertidumbre de medición asociada.
- Explorar la aplicabilidad de la fusión de información para el mejoramiento de incertidumbre en sistemas múltiples mediciones en variables físicas diferentes al flujo. Esta aplicabilidad permitiría el aumento de confiabilidad de información de fuentes no instrumentadas.
- Verificar el desempeño de la metodología en condiciones de fallas simultáneas.

Bibliografía

- [1] LI Y., SUN L., WANG L. & LI ZANG L. Design of Leakage Detection and Location System for Long Range Crude Oil Pipeline, Xiamen, China: 2010 8th IEEE International Conference on Control and Automation, 2010.
- [2] MERIBOUT M., GALEEL M.A., MARZOUQI M.A. & AASI M.A. New Concept for an Effective Leak Detection in Multiphase Fluid Pipelines, United Arab Emirates: First International Conference on Sensor Device Technologies and Applications, 2010.
- [3] AVELINO A., DE PAIVA J. & DA SILVA R. Real Time Leak Detection System Applied to Oil Pipelines Using Sonic Technology and Neural Networks, Natal, Brazil: Federal Institute of Science and Technology of Rio Grande do Norte, 2009.
- [4] FENG J. & HUAGUANG Z. Oil Pipeline Leak Detection and Location Using Double Sensors Pressure Gradient Method, Shenyang, China: Northeastern University, 2004.
- [5] SANTOS A., & YOUNIS M. Sensor Network for Non-Intrusive and Efficient Leak Detection in Long Pipelines, Baltimore: University of Maryland, 2010.
- [6] AMERICAN PETROLEUM INSTITUTE - API. STD 2560 - Reconciliation of Liquid Pipeline Quantities, Washington, D.C.: API Publishing Services, 2005.
- [7] KOKS, D. & CHALLA, S. An Introduction to Bayesian and Dempster-Shafer Data Fusion, Edinburgh, SA: DSTO Systems Sciences Laboratory, 2005.

- [8] AMERICAN PETROLEUM INSTITUTE - API. API MPMS 5.3 - Measurement of Liquid Hydrocarbons by Turbine Meters, Washington, D.C.: API Publishing Services, 2005.
- [9] AMERICAN PETROLEUM INSTITUTE - API. *API MPMS 5.8 - Measurement of Liquid Hydrocarbons by Ultrasonic Flow Meters Using Transit Time Technology*, Washington, D.C.: API Publishing Services, 2005.
- [10] PLANT ENGINEERING MAGAZINE. "Liquid Flowmeters". {En línea}. {11 de agosto de 2011} disponible en: (<http://www.omega.com/techref/flowcontrol.html>)
- [11] GENERAL ELECTRIC COMPANY – GE. "*DigitalFlow™ XMT868i-Panametrics Liquid Flow Ultrasonic Transmitter*" {En línea}. {21 junio de 2005} disponible en: (<http://www.ge-mcs.com/download/co2-flow/920-170C-LR.pdf>)
- [12] ORGANISATION INTERNATIONALE DE MÉTROLOGIE LÉGALE – OIML. OIML G 1-100 - Evaluation of measurement data – Guide to the expression of uncertainty in measurement, Paris: OIML Publications, 2008.
- [13] MAHLER, R. Statistical Multisource-Multitarget Information Fusion, Norwood, MA: Artech House Inc, 2007.
- [14] SCHÖRGENDORFER. A. Extended Confidence-Weighted Averaging in Sensor Fusion, Wien: Vienna University of Technology, 2006.
- [15] ORGANISATION INTERNATIONALE DE MÉTROLOGIE LÉGALE – OIML. OIML G 1-102 -Evaluation of measurement data | Supplement 2 to the Guide to the expression of uncertainty in measurement" | Extension to any number of output quantities, Paris: OIML Publications, 2011.
- [16] W. ELMENREICH. Sensor Fusion in Time-Triggered System, Technischen Universität Wien, Wien, 2002.

-
- [17] D. L. HALL. Mathematical Techniques in Multi-Sensor Data Fusion, Norwood, Massachusetts: Ed Artech House, 1992.
- [18] RODRÍGUEZ M. & GÓMEZ J. Analysis of Three Different Kalman Filter Implementations for Agricultural Vehicle Positioning, Valladolid: The Open Agriculture Journal, 2009.
- [19] TIAN J., GAO M. & LI K. The Detection System of Oil Tube Defect Based on Multisensor Data Fusion by Classify Support Vector Machine, Beijing: First Conference on Innovative Computing, Information and Control, 2006.
- [20] PONG P. & CHALLA S. Empirical Analysis of Generalised Uncertainty Measures with Dempster Shafer Fusion, Melbourne AUS: University of Melbourne, 2007.
- [21] ABDELGAWAD A., MERHI Z., ELGAMEL M. & BAYOUMI M. Multisensor Data Fusion Methods for Petroleum Engineering Applications, Lafayette LA: SAS 2009 - IEEE Sensor Applications Symposium. 2009.
- [22] JIANG Z., CHONQING H., YINGBAO Z. & YANYING W. Research on Crude Oil Pipeline Leakage Detection and Location Based on Information Fusion, Shijiazhuang: First International Workshop on Education Technology and Computer Science, 2009.
- [23] KANGERUD J. Sensor Fusion: Applying sensor fusion in a district heating substation, Ronneby: Blekinge Institute of Technology, 2005.
- [24] ZHONGBO L. Leak Detection by Multi-sensor Fusion Method, Northridge: California State University, 2012.
- [25] TANG. J., GU J. & CAI Z. Data Fusion with Different Accuracy, Shengyang: IEEE International Conference on Robotics and Biomimetics, 2004.
- [26] ZHIZHONG Z., TAO L., YOUHUA W. & HAIPING X. Data Fusion in Vacuum Level Measuring System, Tianjin: Heibi University of Technology, 2009.

- [27] DURRANT-WHYTE H. Multi Sensor Data Fusion, Sydney: The University of Sydney, 2001.
- [28] VERSHINNIN Y. A Data Fusion Algorithm for Multisensor Systems, Coventry: Coventry University, 2002.
- [29] D'AGOSTINI G. A Theory of Measurement Uncertainty Based on Conditional Probability, Rome: Università "La Sapienza" & INFN, 1996.
- [30] QIAN S., STOW C. & BORSUK M. On Monte Carlo methods for Bayesian inference, Durham: Elsevier Science B.V., 2002.
- [31] JOHANNESSON G., HANLEY B. & NITAO J. Dynamic Bayesian Models via Monte Carlo, Livermore CA: Lawrence Livermore National Laboratory, University of California, 2004.
- [32] STEINER G. Model-based Data Fusion in Industrial Process Instrumentation, Wien: Graz University of Technology, 2009.
- [33] SHIH-CHI LO, HWONG-WEN M. & SHANG-LIEN L. Quantifying and reducing uncertainty in life cycle assessment using the Bayesian Monte Carlo method, Taipei: Elsevier Science B.V., 2004.
- [34] AMERICAN PETROLEUM INSTITUTE - API. RP86 - Measurement of Multiphase Flow, Washington, D.C.: API Publishing Services, 2005.
- [35] YANG L., WEN Y., & LI P. Leak Acoustic Detection in Water Distribution Pipelines, Chongqing, China: Proceedings of the 7th World Congress on Intelligent Control and Automation, 2008.

A. Anexo: Simulación en Condiciones de Operación Reales

Este anexo se excluye del documento dado que la información utilizada es de propiedad externa a la Universidad Nacional de Colombia.

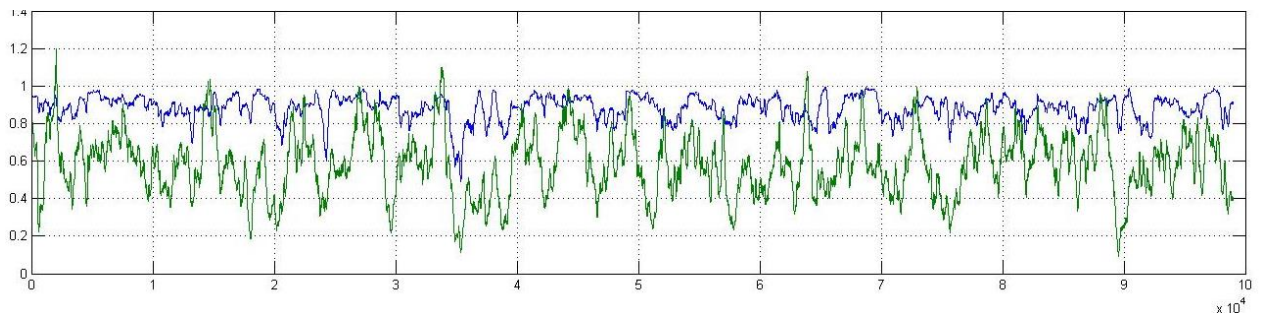
B. Anexo: Umbrales de Sensibilidad

La metodología propuesta anteriormente varía la sensibilidad de detección dependiendo del tamaño de la pérdida o ganancia vista en el segmento de la línea. Para identificar y aislar la desviación es necesario establecer umbrales, los cuales a su vez alertarán y alarmarán a los operadores para tomar acción sobre la operación. Estos umbrales se deben definir en condiciones normales de operación (condición cero) y condiciones desviación operativa (condición L/G).

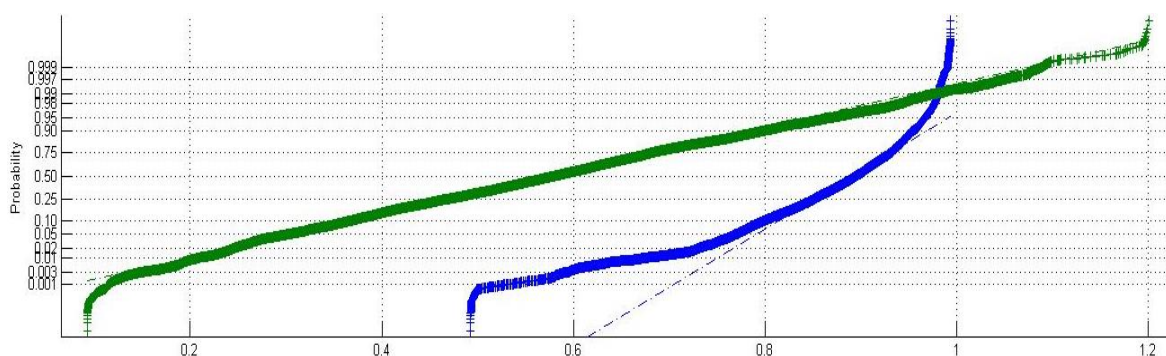
Umbral con Condición Cero

Para definir este umbral, se generará un número considerable de muestras de operación normales (sin desviaciones de ganancia o pérdida) sobre un modelo de la figura 3-4 de la página 53. Posteriormente se aplicará la metodología (Por BY y DS) sobre la información generada y se analizará el comportamiento estadísticamente; se obtendrá el promedio, la desviación estándar de la muestra y su curtosis.

Figura B-1: Definición de umbral con Dempster- Shafer y Bayes²⁴.



²⁴ La tendencia en color verde corresponde a valores utilizando BY y la azul con DS.

Figura B-2: Distribución valores para el umbral cero con DS y BY²⁵.**Tabla B-1:** Valores estadísticos para el análisis de umbral cero con los métodos analizados.

Método	Promedio	Mínimo	Máximo	D.Estándar	Curtosis
Dempster-Shafer	0,8841	0,4925	0,9931	0,0642	6,0443
Bayes	0,5801	0,0926	1,2007	0,1675	2,9364

Observando las figuras anteriores y los valores tabulados se puede resaltar lo siguiente:

- En condiciones de operación normales, la metodología propuesta permite una discriminación vía umbrales fijos por los métodos propuestos (DS y BY).
- La metodología por DS es estable estadísticamente y se satura a un valor cercano al 0,49, todo valor inferior deberá ser una desviación operativa. El caso contrario para los valores tomados con BY, esto debido a que en las muestras tomadas pueden encontrarse en un espectro más amplio y generar falsas alarmas. El umbral para BY se tomará como el promedio menos una desviación estándar, es decir 0,4126 para tener una confiabilidad mayor al 68% de las muestras tomadas.
- Los valores de agrupación de las muestras (curtosis) es más alto para DS que para BY y genera una mayor confiabilidad al uso de la misma.

²⁵ La tendencia en color verde corresponde a valores utilizando BY y Azul con DS

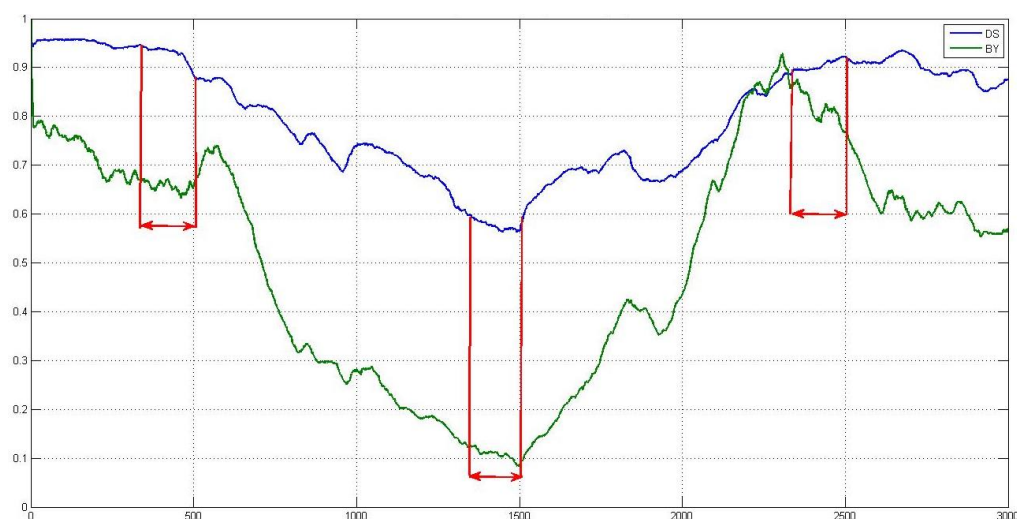
- A pesar de amplio rango de valores que puede tomar en condiciones normales de operación del método Bayes, la homogeneidad de las muestras permiten una mejor detección de desviaciones a pérdidas o ganancias menores al 0,5% del flujo nominal.

Umbral con Desviación de Pérdida o Ganancia

Teniendo definidos los valores para los umbrales de operación normal, se procederá a generar escenarios de pérdida y ganancia donde se demostrará la eficiencia de la metodología para la detección de desviaciones operativas.

Se tomará entonces el modelo simplificado de operación para los escenarios de pérdida o ganancia de la figura 3-4 de la página 53 nuevamente. Se generará escenarios pequeños (4000 muestras) donde se incorporará una pérdida o ganancia. Finalmente se tomarán un número de muestras pequeña (100 muestras) consecutivas y se analizarán estadísticamente (Las muestras se tomaran en los rangos 400:500, 1400:1500 y 2400:2500 como lo muestra en la figura B-3). El valor promedio de este grupo comparado con el umbral fijo de operación normal permitirá generar la alarma según la condición.

Figura B-3: Segmentación para análisis de umbrales de sensibilidad.



Se efectuará una tabulación de la probabilidad de detección de la desviación en proporción al nivel de la desviación que se tendría. Esta tabla permitirá analizar la efectividad del método aplicado y propuesto en esta metodología. La probabilidad de detección se obtendrá de la cantidad de escenarios donde se sobrepasó el umbral de operación normal contrastado con la cantidad de escenarios generados de 10000 escenarios para cada desviación de pérdida o ganancia dado.

Tabla B-2: Valores estadísticos para el análisis de umbral cero con los métodos analizados.

LG	Dempster Shafer			Bayes		
	Sector Inicial	Sector LG	Sector Final	Sector Inicial	Sector LG	Sector Final
1%	0,99	0,95	0,97	0,97	0,96	0,83
0,90%	0,97	0,90	0,95	0,94	0,96	0,82
0,80%	1,00	0,71	0,98	0,97	0,92	0,86
0,70%	0,99	0,63	1,00	0,98	0,86	0,88
0,60%	1,00	0,50	0,96	0,97	0,73	0,89
0,50%	0,98	0,23	0,98	0,96	0,50	0,93
0,40%	0,97	0,18	0,99	0,95	0,46	0,93
0,30%	0,98	0,05	0,98	0,92	0,25	0,88
0,20%	1,00	0,01	1,00	0,93	0,13	0,94
0,10%	0,98	0,01	0,99	0,90	0,14	0,88
0,00%	0,99	0,01	1,00	0,93	0,06	0,88
-0,10%	0,97	0,02	0,98	0,94	0,06	0,88
-0,20%	1,00	0,05	0,99	0,96	0,15	0,92
-0,30%	0,98	0,09	1,00	0,96	0,3	0,92
-0,40%	0,98	0,09	0,98	0,97	0,34	0,93
-0,50%	0,98	0,25	1,00	0,94	0,52	0,91
-0,60%	0,98	0,46	0,99	0,94	0,77	0,93
-0,70%	1,00	0,64	0,96	0,95	0,86	0,87
-0,80%	0,99	0,76	0,98	0,95	0,92	0,89
-0,90%	0,98	0,9	0,94	0,97	0,94	0,86
-1,00%	0,99	0,95	0,98	0,97	0,98	0,86

De la información entregada se puede dar las siguientes observaciones:

- Ambos métodos segmentan la gran mayoría de las desviaciones mayores al 1,0%.
- Ambos métodos segmentan pérdidas o ganancias mayores al 0,8% con una efectividad mayor al 70% de las muestras generadas.
- El método BY permite una segmentación hasta el 0,5% con una efectividad del 50%. El método DS lo haría hasta el 0,6% con 50% de efectividad.

- En caso de no haber una desviación, los métodos reportarían una tasa menor 6% de falsas alarmas. El método DS reportaría una tasa menor a desviaciones menores que el BY, esto es dado a la rata de estabilidad del mismo método.
- Los métodos presentados pueden ser utilizados en conjunto para un análisis amplio de la operación del sistema. Esto es totalmente posible dado que no son mutuamente excluyentes por ser metaprocesos que parten de una fuente única. Esto aumentaría el recurso tecnológico tales como servidores y computadores.

Figura B-4: Resultados de análisis probabilístico de métodos.

